Vol. XXXII. Parts III & IV April, 1942

BIOMETRIKA

A JOURNAL FOR THE STATISTICAL STUDY OF BIOLOGICAL PROBLEMS

FOUNDED BY

W. F. R. WELDON, FRANCIS GALTON AND KARL PEARSON

EDITED BY

EGON S. PEARSON

IN CONSULTATION WITH

HARALD CRAMÉR — J. B. S. HALDANE
R. C. GEARY — G. M. MORANT
MAJOR GREENWOOD — JOHN WISHART

Reprinted by offset-litho, 1952, 1960

ISSUED BY THE BIOMETRIKA OFFICE
UNIVERSITY COLLEGE, LONDON
AND PRINTED AT THE
UNIVERSITY PRESS, CAMBRIDGE

PRINTED IN GREAT BRITAIN

[*Issued* 22 *April* 1942]

VOLUME XXXII. PARTS III AND IV APRIL, 1942

MEDICAL STATISTICS FROM GRAUNT TO FARR
(Continued*)

BY MAJOR GREENWOOD

III. THE STATISTICAL WORK OF GRAUNT

JOHN GRAUNT'S contribution to our subject has always been regarded as one of the great classics of science. A few have indeed doubted whether so great a work could have been achieved by one whose material success was so modest and have sought to transfer the glory to Graunt's highly successful friend Petty. This dispute I relegate to an appendix. I assume that Graunt's published book is substantially his own original work.

The history of the material Graunt used has been written more than once and I have nothing to add to Prof. Hull's story. Graunt had, for a period of more than 60 years, arithmetical statements of the numbers of males and females christened and buried and of the causes of death (not distinguished by sex) under some sixty headings. He had no information as to the ages at death. He had no information as to the number or ages of the living population.

The first act of a scientific statistician is to assess the trustworthiness of his data, criticize his sources. This tedious preliminary to the doing of sums was not much to Petty's taste. Petty, as we have seen, often used different data to reach some conclusion, but hardly ever discusses the reliabilities of the several data. Other Fellows of our College since Petty's day have made the same mistake. The terrible 'howler' committed by Dr William Heberden the younger, and detected, not without satisfaction, by Charles Creighton is classical.† But that was not a unique instance. Indeed, even trained statisticians sometimes confuse names with things. More than one rate of mortality has risen (or fallen) only on paper. Graunt made no such mistakes.

Graunt's *general* argument is that many causes of death are 'but matters of sense', for instance, whether a child were abortive or stillborn, and that in many cases the searchers are 'able to report the opinion of the physician, who was with the patient as they receive the same from the friends of the defunct'. But sometimes the searchers will be wrong and often enough the error will not matter.

> As for consumptions, if the searchers do but truly report (as they may) whether the dead corpse were very lean and worn away, it matters not to many of our purposes whether the diseases were exactly the same, as physicians define it in their books. Moreover, in case a man of seventy-five years old died of a cough (of which had he been free, he might

* The earlier sections were printed in *Biometrika*, **32**, 101–27.

† Creighton, *History of Epidemics in Britain*, **2**, 747–8. Heberden supposed (erroneously) that 'Griping of the Guts' of the Bills was Dysentery and had decreased. It was Infantile Diarrhoea and had simply been transferred to the rubric 'Convulsions'.

have possibly lived to ninety) I esteem it little error (as to many of our purposes) if this person be in the table of casualties, reckoned among the aged, and not placed under the title of coughs (348).*

No doubt this brutal common sense might set on edge the teeth of some Fellows of the College of Physicians even in the seventeenth century, but it was one of the qualities which made Graunt a pioneer. Making the best the enemy of the good is a sure way to hinder any statistical progress. The scientific purist, who will wait for medical statistics until they are nosologically exact, is no wiser than Horace's rustic waiting for the river to flow away.

Graunt, however, did not accept statements which he had the means of testing. Finding in a series of years that of more than a quarter of a million deaths only 392 were assigned to the Pox, he did not infer that Syphilis had been over-rated as a cause of death.

Forasmuch as by the ordinary discourse of the world it seems a great part of men have, at one time or other, had some species of this disease, I wondering why so few died of it, especially because I could not take that to be so harmless, whereof so many complained very fiercely; upon enquiry, I found that those who died of it out of the hospitals (especially that of Kingsland, and the Lock in Southwark) were returned of ulcers and sores. And in brief, I found, that all mentioned to die of the French Pox were returned by the clerks of St Giles' and St Martin's in the Fields only, in which places I understood that most of the vilest and most miserable houses of uncleanness were: from whence I concluded, that only hated persons, and such, whose very noses were eaten off were reported by the searchers to have died of this too frequent malady (356).

In principle, the argument is still valid.

His next example of criticism is the case of Rickets, which first appeared in the Bills of Mortality in 1634 and then with 14 deaths only, but by 1659 had risen to 441. Was Rickets a 'new disease' or did an old disease receive, in the Bills, a new name?

To clear this difficulty out of the bills (for I dare venture on no deeper arguments) I enquired what other casualty before the year 1634, named in the Bills, was most like the rickets; and I found, not only by pretenders to know it, but also from other Bills, that livergrown was the nearest. For in some years I find livergrown, spleen, and rickets, put all together, by reason (as I conceive) of their likeness to each other. Hereupon I added the livergrowns of the year 1634, viz. 77, to the rickets of the same year, viz. 14, making in all 91; which total, as also the number 77 itself, I compared with the livergrowns of the precedent year 1635, viz. 82. All which showed me, that the rickets was a new disease over and above. Now, this being but a faint argument, I looked both forwards and backwards, and found that in the year 1629, when no rickets appeared there were but 94 livergrowns; and in the year 1636 there were 99 livergrowns, although there were also 50 of the rickets: only this is not to be denied, that when the rickets grew very numerous (as in the year 1660, viz. 521) then there appeared not above 15 of livergrown. In the year 1659 were 441 rickets and 8 livergrown; in the year 1658 were 476 rickets and 51 livergrown. Now though it be granted that these diseases were confounded in the judgment of the nurses, yet it is most certain that the livergrown did never but once, viz. anno 1630, exceed 100; whereas anno

* Numbers in brackets are page references to Prof. Hull's edition of *The Economic Writings of Sir William Petty together with the Observations upon the Bills of Mortality more probably by Captain John Graunt*, Cambridge, 1899.

1660, livergrown and rickets were 536. It is also to be observed, that the rickets were never more numerous than now, and that they are still increasing; for anno 1649, there were but 190, next year 260, next after that 329 and so forwards, with some little starting backwards in some years, until the year 1660, which produced the greatest of all (357–8).

This is an excellent statistical argument, and, incidentally, evidence that Graunt wrote his own book, for a physician would probably have suggested that the professional interest excited by the classical treatise of Glisson (assisted by Regimonter) which was published in 1650 might easily have increased the popularity of the diagnosis. Petty, who, with Glisson, was a founder of the Royal Society, would hardly have ignored his colleague's work.

I cannot resist the desire to mention others which, while of little statistical importance, have a medical attraction. Graunt noticed that Stopping of the Stomach first appeared in the Bills of 1636, increased from 6 to 29 by 1647, by 1655 it reached 145, in 1657, 277 and 1660, 314. First he conjectured that Stopping of the Stomach might be the Green Sickness, 'forasmuch as I find few or none to have been returned upon that account, although many be visibly stained with it'. He thought that possibly Green Sickness might not appear in the Bills 'for since the world believes that marriage cures it, it may seem indeed a shame, that any maid should die uncured, when there are more males than females, that is, an overplus of husbands to all that can be wives'. Then he wondered whether Stopping of the Stomach might not be Mother, 'forasmuch I have heard of many troubled with Mother Fits (as they call them) although few returned to have died of them'. But he was diverted by guessing 'rather the Rising of the Lights might be it'. He remembered that some women troubled with the Mother fits did complain of a choking in their throats. 'Now, as I understand, it is more conceivable that the Lights or Lungs (which I have heard called the bellows of the body) not blowing, that is, neither venting out, nor taking in breath, might rather cause such a choking, than that the Mother should rise up thither, and do it. For methinks, when a woman is with child, there is a greater rising, and yet no such fits at all' (359). He notes that Rising of the Lights increased in the Bills from 44 in 1629 to 249 in 1660.

Finally, he suggests a correlation between Stopping of the Stomach, Rising of the Lights in adults and the Livergrown, Spleen and Rickets of children. 'And that what is the Rickets in children, may be the other in more grown bodies; for surely children which recover of the Rickets, may retain somewhat to cause what I have imagined: but of this let the learned physicians consider, as I presume they have' (359).

It might be suggested that one item under Stopping of the Stomach could be surgical, viz. strangulated hernia. Rupture was a heading in the Bills, but the numbers are small and show no regular increase with the increase of population. Graunt's attraction to what used to be called hysterical stigmata is interesting. One wonders how far these passages reflect conversations with Petty. It is clear that Graunt had no belief in the peripatetic uterus; Petty

would have had none. The best medical opinion of the age is, of course, that of Sydenham. Sydenham (whose pathology was traditional) had a pneumatist aetiology of Hysteria, the origin was an ataxia of the animal spirit (which was the pneuma zotikon of ancient tradition). He not only believed that Hysteria might be a serious or even mortal complication of organic disease—as we do still—but that the ataxic spirits might themselves produce humoral corruption and lead to chlorosis or ovarian dropsy (*Dissertatio epistolaris*, 92). So there is nothing repugnant to the best professional opinion of the age in admitting Hysteria to the list of causes of death. Nor is there any gross absurdity in the suggested correlation of increasing Rickets and increasing Hysteria, from the point of view of a layman. But that surmise does not imply any professional hint, it rather suggests a belief in a merely physical factor, the pressure of an enlarged organ. That passage would not have been written by a physician.

These are sufficient instances of Graunt's criticism of sources—the temptation to go on quoting examples must be resisted. I pass to his great achievement, the estimation of rates of mortality at ages when the numbers and ages of the living were not recorded. For such an estimation to be correct, we all know that the population must be stationary, viz. non-increasing, not subject to migration and having constant rates of mortality in the several age groups.

It is a nice point whether Graunt or Petty appreciated the importance of these considerations. Graunt was certainly alive to the fact that the population of London was growing and that the growth was due to immigration from the country. The arithmetical position was this. In the earlier years of his series burials and christenings were about equal in numbers, in 1605 there were 5948 burials and 6504 christenings; in 1625, 7850 burials and 7682 christenings, in 1635, 10,651 burials and 10,034 christenings. Later the burials continued to increase, but the christenings either decreased or failed to increase in the same proportion. This Graunt attributed to neglect of christening owing to religious dissidence and gave excellent reasons for his view. It is clear then that there were two factors of increase, immigration and increasing numbers of births. Most of Graunt's deductions are based upon an analysis of the deaths by causes for twenty years, 1629–36 and 1647–58, which he selected as years comparatively unaffected by plague (of his total of 229,250 deaths only 16,000 were from plague).

If we treat this total as a denominator (or one-twentieth of it) it will, from the point of view of calculating mortality ratios, be affected by two errors. The deaths of immigrants will make it too large and the increasing births will make it too small. Can it be Graunt held that the errors balanced so that, arithmetically speaking, one might behave as if one were dealing with a stationary population? An alternative explanation is that Graunt did not realize the limitations of the method.

A third possibility is that, although he knew the fallacy, he believed that

the incorrect method gave an approximation to truth sufficient for his purposes. This is the solution I should be inclined to adopt were I forced to choose.

As I have pointed out above, there is at least a suggestion that Petty did have some glimmering of the conditions to be fulfilled if a summation of deaths is to give a correct view of rates of mortality. I do not believe that Graunt was less informed on any point of vital statistics than Petty. However, all this is guess-work.

Graunt did not know the ages of the dead; what he did was to pick out of the list of causes of deaths those which he thought lighted only upon children 'not more than four or five years old'. He chose Thrush, Convulsions, Rickets, Teeth and Worms, Abortives, Chrysomes, Infants, Livergrown and Overlaid. These gave him some 70,000 out of some 229,000. Then he assigned half the deaths from Small Pox, Swine Pox, Measles and Worms without Convulsions also to children under six and reaches the final conclusion that about '36% of all quick conceptions die before six years old'.

Is this conclusion—I will not say correct, because we have no data to reach a correct result—but of a reasonable order of magnitude? The answer is that it is eminently reasonable. Two hundred years after Graunt's death, William Farr printed (in the famous Supplement to the *35th Annual Report of the Registrar-General*, p. cxxxvi) an outline Life Table for London. This was, of course, computed by an approximately correct method, using knowledge of the numbers and ages of the living population, and reflects the conditions of seventy-five years ago. Interpolating in this we find that about 32% of 'quick conceptions died before six years old'. There is no good medical reason for holding that the conditions of *child* life in London in the middle of Victoria's reign were much better than in the seventeenth century. The old genius used a bow with a frayed string and made no allowance for windage, but his arrow hit the target not far from the white. He gave the first quantitative measure of the Herodian sacrifice in towns, a sacrifice which was to continue to be offered for more than 200 years.

Graunt then passed to the other end of life and found that 7% of the dead were 'aged'. He conceived that the searchers would mean by 'aged' persons of 70 years or upwards, 'for no man be said to die properly of Age who is much less'. His following suggestion that the proportion living beyond 70 might be used as a measure of healthfulness is not happy. But this calculation may have led him to make, or insert, the most famous passage in his book, viz. what is, in form, the first Life Table ever published.

> Whereas we have found, that of 100 quick Conceptions about 36 of them die before they be six years old, and that perhaps but one surviveth 76; we having seven decads between six and 76, we sought six mean proportional numbers between 64, the remainder, living at six years, and the one, which survives 76, and find that the numbers following are practically near enough to the truth; for men do not die in exact proportions, nor in fractions, from whence arise this Table following (386).

Graunt's figures are 100, 64, 40, 25, 16, 10, 6, 3, 1.

The one survivor to 76 is, as Graunt implies, a guess; perhaps he conjectured that his seven survivors beyond 70 died one a year. How he calculated his mean proportional numbers is unknown. Prof. Willcox conjectured that he experimented with multipliers of 5/8 and 2/3—the former nearly reproduces the figures (see Willcox, *Revue de l'Inst. Intern. de Statistique*, **5** (1937), 327). Ptoukha (*Congrès Intern. de la Population*; Démographie historique, p. 71, Paris, 1937) ingeniously suggests that he used the multiplier $(64-1)/100$ or 0·63.

We must, I fear, conclude sorrowfully that this shot did not find the bull's eye. If Graunt's survivors are compared with those shown in Halley's table (when correctly used, *vide infra*), for 100, 64, 40, 25, 16, 10, 6, 3, 1, we should have 100, 56, 50, 45, 38, 31, 22, 14, 6. It is possible that child mortality was lower in London than in Breslau, but quite incredible that later age mortality should have been so enormously higher.

But, of course, having regard to the data, it would have been more than genius, it would have been magic, had a correct result been obtained.

Prof. Willcox, whose opinion of Graunt is almost as high as mine, regards the passage as inserted on the recommendation of Petty and as Petty's composition. He thinks that it lacks Graunt's caution and suggests the flighty ingenuity of his friend. Prof. Willcox's arguments are weighty, but I am not convinced. That Graunt did not—to use the expressive slang—*feature* his table is true. It is also true (*vide supra*) that passages in Petty's undoubted writings imply that he had some conception of a survivorship table. But—and this is my main difficulty—if this were Petty's idea, I find it difficult to believe that he would not have exploited it. Halley, whose economic scent was not so keen as Petty's, saw the epoch-making importance of an idea which was to transform the business of selling annuities. It would be odd if Petty *had* seen it that he did not comment upon it. Graunt might well have hesitated, being a cautious statistician, but surely not Petty.

However, in spite of modern practice, the writing of history wholly in terms of psychology has its pitfalls.

Let us return to simpler applications of shop arithmetic. The advantages of country life over town life from the point of view of both mortality and morality had been a commonplace of poets, particularly those Roman poets who spent much of their lives in a city, long before the seventeenth century. Graunt was the first to apply an arithmetical test of mortality; he compared the statistics of Romsey with those of London. For Romsey he had ninety years' data of marriage, christenings and burials.

His statements about the population of the parish are not quite consistent. In one sentence he says that it 'both 90 years ago, and also now, consisted of about 2700', but a few lines later says 'it neither appears by the burials, christenings, or by the built of new housing, that the said parish is more populous now, than 90 years ago, by above two or 300 souls'. A little later he says 'it is

clear that the said parish is increased about 300, and it is probable that 3 or four hundred more went to London; and it is known that about 400 went to New England, the Caribe Islands and Newfoundland within these last forty years' (389). Actually, from an estimate of the number of communicants (which he assumes to be rather more than half the total population) he makes the average population between 2700 and 2800. Taking the average of burials for the whole period to be 58, this gives him a death rate of a little more than one in 50, which he contrasts with the London figure of one in 32 (apparently based on his count of 11 families with 88 persons amongst whom 3 deaths occurred in a year; but this is a rate of one in 29).

There is no doubt a certain sketchiness about this, but it was not unreasonable to infer that the Romsey rate was much lower than the London rate.

Graunt found that, unlike London, Romsey had an average excess of christenings over burials, they were in the ratio of 5 to 4. He estimates that over the period the natural increase was 1059, and, as will be seen from the quotation made, he allots about a third of this respectively to London, to the colonies and to the parish itself. He argues that supposing the population of all England to be fourteen times that of London and other parishes to send one-third of their natural increase to London, then the London burials should increase about 200 per annum 'and will answer the increase we observe'.

Here again the argument is reasonable. He goes on to an investigation which has been severely criticized. He gives a table of the greatest and least number of burials in each of the ten-year periods for which he has data. In each decade but one the maximum is more than twice the minimum. But, he remarks, in no decade in the London experience is the largest number of burials twice the smallest number (he excludes deaths from plague from his statistics). 'Which shews, that the opener and freer airs are most subject both to the good and bad impressions, and that the fumes, steams and stenches of London do so medicate and impregnate the air about it, that it becomes capable of little more, as if the said fumes rising out of London met with, opposed and jostled backwards the influences falling from above, or resisted the incursion of the country airs' (392).

Prof. Hull shook his head over this passage. 'This is an attempt to explain by physical conditions the wide range in the observed country death rate which is really due to the narrowness of the field—a single market town—under investigation. It is perhaps the gravest statistical mistake that can be charged against Graunt' (lxxvii).

I do not like to leave a hero in the lurch. I must concede that if both Romsey and London burials were samples from a Poisson universe, the fact that the Poisson parameter for London was at least a hundred times that for Romsey would make it incredible that the London range, in terms of the mean or of the standard deviation, should be so wide as that for Romsey. But Prof. Hull was

wrong in supposing that the wide range in the Romsey rates was due to the narrowness of the field of observation in a statistical sense.

Taking Graunt's 58 as the 'expected' annual deaths then, as 1/58 is small, the Poisson distribution is not far from the symmetry of a normal curve, and using the results of Tippett and E. S. Pearson, we may conclude that the expected range would be $23{\cdot}45 \pm 6{\cdot}073$. The observed ranges for the successive decades are 32, 48, 78, 23, 65, 39, 121, 91, 52. All but one is greater than the expectation and six diverge by more than three times the standard error.

Something more than small numbers is involved. Still, it must be confessed that Graunt did not anticipate the reasoning of James Bernoulli, although an intuition of genius may have led him to think that something more than 'chance' had play here.

Graunt devoted special attention to the demographic influence of the plague. In the first place, he remarks that the attribution plague understated the mortality due to plague. He infers this from the fact that in plague years burials from other causes exceeded the average greatly, 'from whence we may probably suspect, that about 1/4 part more died of the plague than are returned for such'. Next he inferred that after a great outburst plague lingered for several years.

> The plague of 1636 lasted twelve years, in eight whereof there died 2000 per annum one with another, and never under 300. The which shows that the contagion of the plague depends more upon the disposition of the air than upon the effluvia from the bodies of men. Which also we prove by the sudden jumps which the plague hath made, leaping in one week from 118 to 927; and back again from 993 to 258; and from thence again the very next week to 852. The which effects must surely be rather attributed to change of the air, than of the constitutions of men's bodies, otherwise than as this depends upon that (366).

Finally, he observes that within two years the city was re-peopled; a deduction from the time taken for the number of christenings to reach again the level of a pre-plague year.

We may, if we please, smile at Graunt's epidemiological inference. But it is a reasonable inference from the facts when we remember that in Graunt's day—in spite of Fracastorius—contagium was not thought of as contagium vivum, but as a mere sympathetic vibration or passing on of something.

I have, I hope, given an adequate sample of Graunt's quality, but have not mentioned the most famous of all his deductions. Both in London and the country, on the average more males were christened than females, but more males died young or entered celibate occupations. So we reach this conclusion:

> We have hitherto said, there are more males than females; we say next that the one exceed the other by about a thirteenth part. So that although more men die violent deaths than women, that is, more are slain in wars, killed by mischance, drowned at sea and die by the hand of justice; moreover more men go to colonies and travel into foreign parts than women; and lastly, more remain unmarried than of women as fellows of colleges, and apprentices above eighteen etc. yet the said thirteenth part difference bringeth the business but to such a pass, that every woman may have an husband, without the allowance of polygamy (375).

The story of how this arithmetical justification of God's providence attracted the attention of Derham, of how Derham's book fired the enthusiasm of the Prussian Army chaplain Johann Peter Süssmilch and of how Süssmilch's book influenced Malthus has been well told by Hull. I should not myself rank this section high among Graunt's researches. From a demographic point of view neither judicial hangings nor college fellowships could have had much effect in reducing the male excess.

Even copious quotation fails to convey the spirit of a complete book. I have quoted good things, but many more remain. Graunt revealed sundry important truths and not the least important was that very imperfect data, if patiently considered, will tell us something it is good for us to know. If young medical officers going to parts of the empire where organized medical and demographical information is at no higher a level than that of seventeenth-century England—and there are many such places—were restricted to a single book on statistics, I should advise them to take not a modern scientific work, but old John Graunt's *Observations*.

APPENDIX

Did Graunt write the book published over his name?

John Graunt and William Petty were, as we have seen, close friends. Graunt, as the world judges success, failed and Petty succeeded. But by the judgment of scientific men in the seventeenth century, and ever since, the order of intellectual precedence was reversed. From the moment of publication a few discerning people perceived the originality and importance of the *Observations*, the same people who, while admiring Petty's verve, ingenuity and worldly success, did not take over-seriously his bright ideas.

But Graunt was a man of one book. Save a note upon the multiplication of carp and the growth of salmon, he published nothing more. Petty went on writing, scheming and talking for thirteen years after Graunt's death. That often enough in that period, the *Observations* were discussed over the wine—as they were quoted in Petty's writings—we may suppose. That Graunt discussed his work with Petty both before and after publication we may also take for certain, although we have no formal proof of it. The country statistics which Graunt first used were from Petty's native parish and even if we are not disposed—as certainly I am not—to give much weight to particular turns of phraseology, still there are sufficient verbal oddities in some pages of Graunt's book to suggest Petty's hand.

In these circumstances, it would not be very surprising if Petty's associates, particularly those who were not good judges of statistical work, were to conclude that Petty's share in the remarkable achievement of Graunt were greater than

appeared. It is not even judging Petty too harshly to suppose that he himself might come to share the opinion. There is no evidence that Petty ever *did* explicitly claim the credit. In one list of his writings (one of four), found among Petty's Papers, he did include the *Observations*, which at least is evidence that he thought himself entitled to a share of the credit. We may suppose that if, in familiar intercourse, somebody had said 'Come, confess Sir William, yours was the hand that guided the pen of poor John Graunt', he might not have denied it very strenuously. I think I have produced evidence enough that Petty did not under-rate his powers and was not conspicuous for delicacy of feeling. My guess would be that long before his death he did come to believe that Graunt's intellectual success was due to his help.

Whether Petty believed this or not, it is certain that friends and associates of Petty began to believe it soon after Graunt's death, and the belief has been entertained by a few people in each generation since. These, with one conspicuous exception, have been drawn from Petty's friends or descendants or from literary critics.

In the seventeenth century, of Petty's circle, Evelyn, Southwell and Aubrey believed or said that Petty wrote or inspired Graunt's book. Two Fellows of the Royal Society, Houghton and Halley, also attributed the book to Petty. The only one of the five who was certainly a competent judge of scientific merit was Halley. Halley began the memoir which contains his Breslau table with these words:

> The contemplation of the mortality of mankind has, besides the moral, its physical and political uses, both which have some time since been most judiciously considered by the curious Sir William Petty, in his moral and political Observations upon the Bills of Mortality of London, owned by Captain John Graunt. And since in a like treatise on the Bills of Mortality of Dublin.... But the deductions from those bills of mortality seemed even to their authors (*sic*) to be defective. (*Phil. Trans.* no. 196 (1693), p. 596.)

Since the seventeenth century, there has been unanimity among demographic statisticians and economists that Petty could not have written Graunt's book. Halley was quite as good a judge of scientific merit as any of them and a contemporary of the canvassed writers; if I were sure that he had read and compared Petty's acknowledged works with the *Observations* I should prefer his opinion to that of other 'experts'—including, of course, my own. Halley's direct testimony, in the sense of a court of law, would be valueless; he was only six years old when the *Observations* were published and became a Fellow of the Royal Society five years after the death of Graunt. There is no evidence that either before or after the period of writing and publishing his famous memoir, Halley worked on demography. After his memoir, but in his lifetime, a new epoch in mathematical vital statistics began. De Moivre, eleven years younger than Halley, brought out his principal works in the lifetime of Halley (1656–1742) and used Halley's table. The two men must have been well acquainted,

for both were enthusiastic disciples and intimate friends of Newton, but Halley, like Graunt, made only one contribution to the literature of demography.

So it may be doubted whether Halley were sufficiently interested in demographic or economic writings to have read Petty's tracts at all. Also in the passage cited above (apart from the writing of 'authors' not 'author') the collocation of the *Observations* on the Dublin Bills with those on the London Bills is curious. There is no doubt that the *Observations* on the Dublin Bills were the work of Petty, and in the first edition of them they are stated on the title-page to be by the 'Observator on the London Bills of Mortality'. But this, as Prof. Hull pointed out (xlii), was probably a catch-penny device of the publisher, Mark Pardoe, to draw a public which had just taken a fifth edition of the London *Observations*. Actually the book did not sell, and when the publisher reissued an enlarged version, Petty's name appeared on the title-page without any reference to the London *Observations*. I conclude that Halley's evidence is less weighty than it seemed. He may well have had before him copies of Graunt's book and of the two editions of the Dublin *Observations*. Having no other knowledge of the literature he would naturally enough write as he did.

If we eliminate Halley, no other expert countenanced Petty's authorship and one, Augustus de Morgan, gave an amusing but quite cogent reason for dismissing the notion.

In speaking of the variations in the annual numbers of deaths attributed to Rickets, Graunt said:

> Now, such back-starting seem to be universal in all things: for we do not only see in the progressive motion of wheels of watches, and in the rowing of boats, that there is a little starting or jerking backwards between every step forwards, but also (if I am not much deceived) there appeared the like in the motion of the moon, which in the long telescopes at Gresham College one may sensibly discern (358).

De Morgan (*Budget of Paradoxes*, 68; *Assurance Magazine*, **8**, 167) commented on the improbability that 'that excellent machinist, Sir William Petty, who passed his day among the astronomers' would attribute to the motion of the moon in her orbit all the tremors which she gets from a shaky telescope.

Down to 1927 the matter was regarded, in scientific circles, as settled. In that year the late Marquis of Lansdowne published a copious selection of the Petty Papers with what he regarded as new evidence in favour of Petty.

The only new evidence of a direct kind was a manuscript list in Petty's hand of his writings or projected writings which included the *Observations*. There are three other lists which do not include the *Observations*, and if we are to suppose that the entry really referred to the book published under Graunt's name, then we must believe that in 1685 and in 1686 Petty had forgotten his best title to scientific immortality. The remainder of the evidence consists of parallel passages and *ad captandum* arguments to the effect that it was more probable that a physician had written on questions of medical statistics than a tradesman. This

publication led to a lively controversy. Of the merits of this, I, as a party to it, am not an impartial judge. Purely literary arguments do not appeal to me when the question is of scientific method. Thus, Dr L. F. Powell attached weight to the fact that Dr Johnson in conversation had attributed to Petty an observation (not statistical) which is made not in Petty's writings but in Graunt's book.

In the discussion the word 'style' is used in different senses by the combatants. The statisticians are thinking of scientific method, the literary critics of verbal arrangement. To the former the fact that, particularly in the conclusions and the Appendix, Graunt's book has turns of phraseology which suggest Petty's hand, seems of little importance. To the latter it seems very significant.

In the article by Prof. Willcox, which I have quoted above, the controversy is reviewed, and the author concurs generally with his statistical predecessors.

Prof. Willcox does, however, differ from his predecessors in one important particular. He holds that the famous life table was supplied by Petty. He argues that this is far too conjectural to have been the work of so cautious a reasoner as Graunt:

> In attempting to reconstruct its origin I have surmised that after Graunt had estimated that 36 per cent. of the deaths were due to children's diseases, that they all occurred under the age of six, and that the seven per cent. who were reported to have died 'aged' died at over 70 years of age (at one place he says over sixty), he felt unable to go further and reported his difficulty to Petty, already perhaps speculating about a series of similar problems.
>
> Petty guessed at the number of survivors at the end of each decennial age period, 6–15, 16–25 etc. incidentally and characteristically ignoring Graunt's theory that seven per cent. survived seventy, and assuming instead, without reason, that one per cent. survived seventy-six and not one per cent. eighty-six, and that the survivors at age six decreased with each age period in a geometrical progression approximately equal to the 64 per cent. which Graunt had set for the first group (326–7).

Prof. Willcox's argument is cogent. It may be strengthened by a criticism of the late Prof. Westergaard (*Contributions to the History of Statistics* (London, 1932), p. 23). In using this table, Graunt made a serious blunder. In order to estimate the number of men of military age in London he subtracted the number alive at age 56 from the number alive at age 16. But this simply gives him the number *dying* between those ages; what he wanted was some average of the l_x's. It is evident that Graunt was not at all clear in his mind as to how to use a life table.

On the other hand, if this table were really Petty's idea, it is hard to understand why he did not exploit it. If Petty had been a Halley, the explanation would be obvious. The table is wrong; the conditions for the validity of the method were not fulfilled. There is indeed (*vide supra*) some evidence that Petty did know what data were necessary in order to construct a proper life table. One seems on the horns of a dilemma. If Petty thought the table was correct why did he make no further use of the method? If he thought it was wrong, would he have urged Graunt to insert it?

Although Prof. Willcox has certainly shaken my previous conviction, I still feel reluctant to surrender Graunt's table to Petty. However, there may be an element of sentimentality in this. At least the statisticians agree that the answer to the question which I have placed at the head of this Appendix is emphatically *yes*.

REFERENCES TO THE RECENT CONTROVERSY IN CHRONOLOGICAL ORDER

1927. LANSDOWNE, Marquis of. *The Petty Papers.*
1928. GREENWOOD, M. *J.R. Statist. Soc.* **91**, 79.
1928. LANSDOWNE, Marquis of. *The Petty Southwell Correspondence*, pp. xxiii–xxxii.
1932. LANSDOWNE, Marquis of. *The Times Literary Supplement*, 8 Sept.
1932. BRETT-JAMES, N. G. *The Times Literary Supplement*, 15 Sept.
1932. GREENWOOD, M. *The Times Literary Supplement*, 22 Sept.
1932. LANSDOWNE, Marquis of. *The Times Literary Supplement*, 13 Oct.
1932. POWELL, L. F. *The Times Literary Supplement*, 20 Oct.
1933. GREENWOOD, M. *J. R. Statist. Soc.* **96**, 76.
1937. WILLCOX, W. F. *Revue de l'Inst. Intern. de Statistique*, **5**, 321.

IV. HALLEY'S LIFE TABLE

The long and fruitful life of Edmund Halley (1656–1742) belongs to the general history of science; of him it may indeed be said *nihil quod tetigit non ornavit*. He made only one contribution to our subject, but it was of first-rate importance.

The circumstances of this undertaking are obscure; Halley would have perceived the imperfections of Graunt's life table, but it is not known whether it was he who set on foot a search for better statistical material than Graunt had had. Inquiries were however made, and made after he had become a Fellow of the Royal Society, so it is at least possible that Halley, who had travelled extensively in Europe (he was at Danzig in 1679 and in Italy in 1681), suggested that something might be found abroad. By 1691, the King's Librarian, Henry Justel, who was in touch with the Society, had been brought into communication, possibly through Leibniz, with Caspar Neumann, a scientifically minded evangelical pastor of Breslau. Neumann supplied the data which Halley used.

In 1883, J. Graetzer, a medical-statistical official of Breslau, published a little monograph* which throws light upon the work. He not only extracted from the Breslau archives all the data which were or might have been communicated to the Royal Society but had the Society's archives searched, with the result that a letter from Neumann to Justel and another from Neumann to Halley, both with statistical appendices, were discovered. Thanks to the labours of Graetzer and an essay by R. Böckh (*Bulletin de l'Inst. Intern. de Statistique*, **7** (1893), 1) we can form a reasonably clear idea of Halley's method, which was not what those who have not examined the literature suppose it to have been.

It is often stated that Halley, having found that during the five years of observation the number of births only slightly exceeded the number of burials,

* Graetzer, J., *Edmund Halley und Caspar Neumann*, 1883.

treated the population as stationary and constructed a life table by a simple summation of the deaths in the manner already explained. He was much wiser. What he tried to do was to construct a *population* table, in the following way. Suppose we know how many children were born in a calendar year, say 1690, in a town not subject to migration which maintains accurate registers of ages at death, and then we discover how many of the children born in 1690 will be alive on each successive first of January by a series of subtractions. We shall have the survivors on 1 January 1691 by subtracting those of the children born in 1690 who died in 1690. We shall have the survivors on 1 January 1692 by subtracting the deaths in 1691 occurring among the survivors to 1 January 1691, and so on. This will give a precise enumeration of the living population, and this is what Halley wanted. The figures we shall obtain will *not* be the conventional l_x's of a Life or (as German writers say) Mortality Table, but what in most modern books are represented by the capital letter L or years of life lived or "persons" living between the termini (see Appendix). *If* the population is stationary, the sum of these figures gives the population of the place under study. Now for ages between 1 (last birthday) and very advanced ages, L_x is simply l_x diminished by $\frac{1}{2}(l_x - l_{x+1})$. In the first year of life (and at advanced ages) the difference is greater. Thus in the first year of life deaths are not evenly distributed throughout the year of life, more than 70% of them occur in the first six months of life, so that instead of subtracting half the deaths we must subtract nearly three-quarters. Halley himself assigned 68% of deaths in the first year of life to the first half of the first year. The reason why Halley proceeded in this way was that he knew the population *not* to be stationary. His idea was to obtain the figures for the first few years of life accurately—indeed just as they are now obtained—and then to correct for excess of births over deaths.

His masterly plan was partly defeated by the fact that his Breslau correspondent Neumann was not so good a statistician as he was. Halley's letter to Neumann has not been preserved, we have only Neumann's answer of 1 March 1694. Probably Halley asked Neumann to send him (as a check on the calculations he had already made) the exact numbers of survivors on 1 January for five years, of births in a calendar year. Neumann did send him a table, but the table, as Graetzer pointed out, is wrong. Neumann gave the correct figures for 1 January of the first successive year, incorrect figures for the other years. Between 1 January of the year following the year the births of which are under study, and the next first of January, some of those born in the starting year will die under and some over a year of age. Neumann merely deducted the former, so he has too many survivors. To reach the right figure would have meant taking more trouble and he did not appreciate the importance of this. Böckh—whose opinion of his statistical contemporaries has a tinge of bitterness rare, of course, in other scientific pursuits—remarks that it was not strange Neumann should miss the point as it had been missed by many statisticians long after his time.

It is at least clear that Halley had realized an important truth which did not become part of even expert knowledge for more than a century.

The precise arithmetical details of Halley's work are not perhaps of much medical interest. Graetzer and Böckh have done a good deal to clear it up. The data used (an average of five years) had 1238 births and 1174 deaths and the table accounts for 1238 deaths. Halley must therefore have had a plan for increasing deaths. It is likely, from an observation he makes on mortality in Christ's Hospital, that he did not *wholly* depend on the Breslau figures. Graetzer suggests that Halley may have made two graphs, one having an ordinate of 1238 at the origin and an ordinate of 64 at the oldest age, the other an ordinate of 1174 at the origin and 0 at the oldest age and that he plotted the survivors for each graph based on recorded deaths and drew a curve passing through 1238 and 0 between these graphs. It may be so. Using the original material which Graetzer published, Böckh recalculated the table. The results do not, except at ages over 60, differ materially from Halley's. So far as concerns the mean after life time (expectation of life), Halley's table gives 27·54 years at birth, Graetzer's material 27·69. For ages under 40 the re-working gives slightly lower and for later ages higher mortality. It may be noted that Halley's table gives appreciably *higher* mortality in childhood than Graunt's, more than 43% instead of 36% are dead by the age of six years. But Graunt's method would exaggerate mortality (so would Halley's method, but, owing to his precautions, not so greatly). On the other hand, Graunt's estimate of age is only an intelligent guess. Actually, as Graetzer showed, the infant and child mortality shown by Halley's table differed little from the observed rates of mortality in the city of Breslau in 1876–80.

It has been said that Halley was not greatly interested in the medical aspects of his work. After describing methods of calculating the prices of annuities, he has the following passage*:

> It may be objected that the different salubrity of places does hinder this proposal from being universal; nor can it be denied. But by the number that die being 1,174 per annum in 34,000 it does appear that about a 30th. part die yearly as Sir William Petty has computed for London; and the number that die in infancy is a good argument that the air is but indifferently salubrious. So that by what I can learn, there cannot perhaps be one better place proposed for a standard. At least 'tis desired, that in imitation hereof the curious in other cities would attempt something of the same nature, than which nothing perhaps can be more useful.

That the mortality of childhood depends upon the atmosphere is not so foolish an hypothesis as it may seem to us. Halley lived before breast-feeding became the exception rather than the rule. The 'curious' in other cities had not the wit to follow his advice. He made no other contribution to the science of vital statistics; a gain to astronomy but a heavy loss to demography.

* I have read Halley's paper in the collection of papers, many by him, collected under the title *Miscellanea Curiosa*, printed in London in 1705; the quotation is from p. 300.

Appendix

Halley's table is printed in two columns, the first headed 'Age current', the second 'Persons'. Thus:

Age current	Persons
1	1000
2	855
3	798
4	760
5	732
6	710
7	692
8	680
9	670
10	661

and so on.

A mistake sometimes made is to suppose that Halley meant by age current simply the end of each year of life and that the entry against each 'age current' is the number of survivors at exact age one year less than the 'age current', viz. that of 1000 born 855 survived to the first anniversary, 798 to the second anniversary, etc. The fact that Halley uses the round number 1000 for a first entry does something to encourage the mistake among readers who have not consulted the original paper and it is sometimes made by people who should know better. It is actually a terrible 'howler', leading to a wholly false view of rates of mortality in early life. Thus if 1000 and 855 were really the first two entries of a Life Table as set out now, then, as the first two entries in English Life Table no. 7 Males (mortality 1901–10) are 1000 and 856, we might conclude that mortality in the first year of life was no lower in 1901–10 than in Breslau in the last years of the seventeenth century. But the 1000 of Halley's table is *not* the number of new-born children but the average number out of 1238 born living between the ages of 0 and 1. This is what is called the L_x of a modern table or the population living between the ages x and $x+1$. If we have a column of L_x's, which is what Halley gives us, we can deduce therefrom the more familiar l_x's provided we know the starting value and the number of deaths in the first year of life. Halley gives both items. He says that of 1238 annual births 348 die annually. So that his l_0 is not 1000 but 1238, and his l_1 is 890. He chose 1000 for L_0 by assuming that of the 348 deaths in the first year of life 238 occurred in the first six months of life, 68%. This differs very little from the modern practice; in Life Table no. 7 quoted above 73·5% of the deaths in the first year of life are assigned to the first six months of life. Having been given l_0 and l_1 we can deduce the other l's from the values of the L's which Halley gives because, after the end of the first year of life there is little error in supposing that the deaths between two birthdays are evenly distributed over the year; so, for instance, l_2 will be equal to L_1 less half the difference between l_1 and l_2,

and proceeding in this way we put Halley's table into modern form. I attach a table calculated by Böckh.

It will be seen that, if Halley's table is properly used, the comparison is not of 1000 and 855 with 1000 and 856 but of 1000 and 719 with 1000 and 856.

Actually this is still slightly optimistic, because I am comparing 'persons' with males. The 'persons' figure for 1901–10 is 1000, 869. On the other hand, in the Breslau data still births (or some of them) are included in births, so that the mortality is slightly exaggerated. If for instance 7% are still born, the survivors to 1 will be the same, but the l_0 should be reduced to 930. Or alternatively we should write 1000 and 773.

I attach Halley's table reduced to modern form and with the corresponding expectations of life calculated by Böckh (I have reworked some of the values from the data and agree with Böckh's figures).

Halley's Table, expressed in modern form together with the Expectations of Life at quinquennial intervals (Böckh)

Age	l_x	$\mathring{e}_x$	Age	l_x	$\mathring{e}_x$
0	10,000	27·54	40	3,557	22·05
5	5,816	41·47	45	3,167	19·47
10	5,307	40·25	50	2,751	17·05
15	5,049	37·19	55	2,319	14·75
20	4,806	33·93	60	1,914	12·33
25	4,552	30·69	65	1,511	9·96
30	4,257	27·64	70	1,103	7·74
35	3,921	24·78	75	670	7·50

V. GUESSING THE POPULATION

My object is to trace the growth in our country of that part of statistical science which is of interest to students of medicine or public health. In speaking of such pioneers as Graunt, Petty and Halley it was proper to construe the obligation rather freely. Both Graunt and Petty did clearly perceive the relevance of their researches to matters of public health or even clinical medicine, but much of what Petty did had a more direct bearing upon political questions than those of public health. Again, the life table is a way of expressing the facts of mortality which is valuable in some medical researches, but its importance as a statistical instrument has been much greater in non-medical than medical circles, above all of course in the financing of assurance business. The commercial importance of life tables was perceived by Halley and by other mathematicians of his and the following generations.

Looking at the position after Halley's publication it was clear that progress might be made (1) in improving the accuracy of the life table, viz. by obtaining data more relevant to the conditions of life of persons who assured their lives or bought annuities, (2) in simplifying the very laborious calculations which the

determination of praemia or purchase values required. Under (1) no progress worth speaking of was made in England until the end of the eighteenth century. This was partly due, as we shall see, to a not entirely unjustified disbelief in the powers of the medical profession to change the rate of mortality, partly to ignorance. No first-rate English mathematician after Halley gave any critical attention to the theory of the Life Table before the nineteenth century. Under (2) considerable progress was made, but this progress is of little or no medical interest and to describe it would involve entering upon tedious arithmetical and algebraical detail. The primary medical-statistical quaesita are correct enumerations of deaths by sex at ages and by causes, and of the numbers living in sex and age groups. When these have been satisfied, the medical statistician can get to work.

For 150 years after Graunt's death very little was done to improve matters. Down to 1801 the population as a whole had not been counted; forty years more passed before a reasonable age distribution was secured, and it was thirty-eight years after the first denominator (populations) that the first numerator (deaths) of the fundamental fractions was obtained. Until 1801 intelligent guessing was the method and the guesses of the eighteenth century deserve a few pages, if only because they prove that statistical ability is as rare as other kinds of ability and that wishful thinking is not a modern foible.

The first estimator to mention belongs to the seventeenth century and was a younger contemporary of Graunt and Petty, Gregory King (1648–1712). He was born in Lichfield, the son of a land surveyor. At the age of fourteen he was recommended as a clerk to the famous herald Dugdale with whom he worked for several years; after Dugdale had finished his Visitation, King worked for various amateur antiquaries and was eventually invited by a lady of property in Sandon (Staffordshire) to be her steward, auditor and secretary. Here he remained until 1672 when he moved to London and, no doubt through Dugdale's recommendation, had a considerable amount of employment in both heraldic work and ordinary surveying. In 1677 he became a member of the College of Arms, in which he attained the rank of Lancaster Herald and so continued until his death, but worked for other official bodies on financial subjects.

The decorous memoir by George Chalmers, from which I have extracted these particulars, does not give us a very life-like picture of the man himself. There is a certain likeness between the early careers of Petty and King. King was not indeed shipped as a cabin boy, but Mr King (the elder) drank (if we may venture so coarse an abbreviation of Chalmers's statement that the father studied and practised his profession 'with more attention to good fellowship than mathematical studies generally allow') and King junior was a pupil teacher at eleven. If he really read Hesiod and Homer, made Greek verses and taught himself to survey land in his thirteenth year he must have had Petty's precocity. Both Petty and King had experience of practical surveying and, of course, both

were interested in political arithmetic. But there the parallel ends. King was a professional surveyor and archivist and had a reasonably successful professional career. Petty was—Petty. One might, perhaps, adduce as another parallel that King made some enemies and thought himself ill-used. But the job by which Sir John Vanbrugh, a stranger to the College of Arms, was made a king-at-arms over the head of an official of twenty years' standing would have galled the meekest of mankind. One may safely conclude that King had more knowledge of the data of political arithmetic than Petty and less originality. His vital statistical work was not published until nearly a century after his death, as an appendix to the second edition of *An Estimate of the Comparative Strength of Great Britain* by George Chalmers, London, 1803. Perhaps he never intended to publish it—he communicated the substance to his contemporary Davenant—and this may explain why there are no details of how some of his results were reached. The report reads rather like a document prepared for official use by persons interested in results not methods.

The starting-point of King's attempt to estimate the population was a return from the Hearth Office of the number of houses assessed to tax on Lady Day, 1690. That was 1,319,215 which, King estimated, had increased to 1,326,000 by 1695. He deducted 30,000 for empty divided houses,* took the round figure of 1,300,000 and assigned 105,000 to the London area, 195,000 to other cities and market towns and 1,000,000 to villages and hamlets. He used a series of multipliers, 5·4 for a house within the walls of London, 4·6 for a house within the liberties, 4·4 for the out parishes in Surrey and Middlesex and 4·3 for Westminster. For other towns, his multiplier was 4·3 and for villages 4·0.

Having performed his multiplications he gives London a bonus of 10%, other towns 2% and villages 1%. Lastly he estimates homeless people to number 80,000. The final result to the nearest round number is $5\frac{1}{2}$ millions.

How King obtained his multiplier is not clear. In addition to the Hearth Office data he says he used 'the assessments on marriages, births, and burials, parish registers and other public accounts' and that from these he deduced the multipliers, but this is rather vague. He also classified the population by sex, civil state and age (under 1, under 5, under 10, under 16, above 16, above 21, above 25, above 60). How he reached these figures is not explained.

But nothing succeeds like success. As we shall see, his estimate of the total

* Prof. E. C. K. Gonner (*J.R. Statist. Soc.* **76** (1912–13), 261–97), in an interesting paper which I have largely used in writing this chapter, remarks that the 'houses' of the Hearth Office must have been really families or separate occupations as King indeed realized, and thinks that King fell into some confusion in attempting to replace families by houses. Gonner argued that the best way was to proceed on the basis that the Hearth Office unit of a family should be retained and be corrected for empty houses, blacksmiths' shops, etc. on the basis of 1801 census returns and the multiplier used should be persons per family of 1801. The result is to give a figure about a quarter of a million larger than King's. The method described in the text also leads to the conclusion that King somewhat understated the population.

population is probably very near the truth and Prof. Westergaard has remarked that, judging from Swedish observations of a few years later, King's age distribution is quite reasonable.

As a statistical prophet King was no more successful than his contemporaries (and successors). He believed that down to his time the population of England had doubled in 435 years, that the next doubling would require from 1200 to 1300 years and that in A.D. 3500–3600 the population would reach 22 millions of souls, in case, as he cautiously adds 'the world should last so long'. His estimates as a matter of arithmetical curiosity are excellently fitted by a logistic with an upper asymptote of fifteen millions and would give the present population as about eight millions.

Modern statisticians, such as Farr and Brownlee, have confirmed King's estimate of the population at the end of the seventeenth century in the following way. After 1801 the population was known by actual counting and for the first forty years of the nineteenth century baptisms and burials were still the only data of births and deaths. If one started from, say, the enumeration of 1831 and worked back to the population of 1821 by adding the numbers of burials and subtracting the number of baptisms then, if these really measured deaths and births, the result ought to agree with the census enumeration, provided immigration and emigration balanced. But the burials and baptisms understated deaths and births. One might adjust the figures by multipliers to bring the result into agreement with the census and then test against another backward run of ten years. Brownlee found that if the number of burials were multiplied by 1·2 and the number of baptisms by 1·243, the agreement was good.

This may seem a highly conjectural method, but it certainly gives quite good results. The difference between births and deaths estimated in this way for the decennium 1801–10, I find to be about 12·4 per 1000 living. If one multiplies the enumerated population of 1801 by $(1{\cdot}0124)^{10}$ we reach 10·1 millions, not a bad approximation to 10·2 millions actually counted. Assuming that before 1801 burials and baptisms had the same relation to deaths and births as between 1801 and 1841, we can work backwards to the beginning of the eighteenth century with the result that the population then was about 5·8 millions, not much more than King's estimate. In view of the following discussion it will be useful to consider the probable state of the population (as determined by these methods) in the eighteenth century. In the first sixty years of the century it grew very slowly, was about 6·1 millions in 1751 and 6·5 millions in 1761. It then began to increase faster, was 7·5 in 1781, 8·2 by 1791 and 9·2 at the census of 1801 (8·9 as enumerated, but an estimate of a deficit of 1/30th was made).

From Gregory King's time to the census of 1801 we have a series of more or less intelligent guesses.

These are well described in Prof. Gonner's paper.* Two schools of thought

* *J.R. Statist. Soc.* **76** (1912–13), 261–96.

did battle in the eighteenth century; the pessimists who held that the population was decreasing and the country going steadily to the dogs, and the optimists who believed just the contrary. Both used the same weapons. The heavy artillery was a return of houses for taxation purposes increased conjecturally by a figure for houses which escaped taxation, the sum multiplied by a conjectural average of persons per house. As light artillery one had the yield of taxes on commodities and the returns of baptisms and burials.

A pessimist put the number of untaxed houses low and the multiplier low, and an optimist raised both.

The first controversy which took place in 1754 in the proceedings of the Royal Society did not attract much notice. Brackenridge (mildly pessimistic) pointed out that the number of houses assessed to house tax had decreased between 1710 and 1754 from 729,048 to 690,000, which suggested a decrease of population (by a previous conjectural calculation based on burials and baptisms, he had reckoned a small increase, which was probably correct). Much turned on the number of houses which did not pay tax (either because the occupant was in receipt of alms, did not, owing to poverty, contribute to the church or poor rate, or through mere default). Brackenridge put the number at 200,000. His critic, Forster, argued that Brackenridge under-stated the number of untaxed houses, adducing a sample of nine country parishes with 588 houses of which only 177 were taxed and a market town with 229 taxed houses out of 448. Using these figures as a basis for conjecture Forster raises Brackenridge's 890,000 to 1,427,110. From this (with a multiplier of 6 for town houses and 5 for country houses) he reaches a population of seven and a half millions—probably a considerable over-estimate.

The next controversy was a quarter of a century later (in a period when the population was certainly increasing) and its originator was Dr Richard Price (1723–91), who has attained a posthumous celebrity reminiscent of the man whose title to distinction was that he had once been kicked by George IV. Most readers know him as the preacher of a sermon which was the text of Burke's *Reflections*, most students of economic history know him as the inventor of that theory of the virtue of a Sinking Fund which has been likened to the economic system of a community which prospered by taking in one another's washing; most vital statisticians remember him as the computor of the Northampton Life Table which gave a seriously incorrect picture of prevailing mortality and indirectly cost the country a large sum of money. Finally, in the controversy about to be described, Price was pertinaciously in the wrong on all the main issues.

The apparent inference from all this is that Price was either a fool or a knave. Gainsborough's portrait of the Rev. Richard Price, which hangs (or did hang) in the Board Room of the Equitable Assurance Society, gives no support to the hypothesis that Price was a fool; his life would be a promising field of research for a young historian with a competent knowledge of economics. His importance

in statistical history is not great enough to justify me in a critical study (even if I had the necessary training in finance and economics). My guess is that Price was an able, self-confident, original-minded man, who knew a good deal about many things and had no exact knowledge of anything. He had 'a way' with him, he could *interest* people. In fact he had some of the qualities of Petty. It is easy enough to make jokes about his notion of the mysterious power of money to increase at compound interest and it is possible that William Pitt the younger (who was only a boy when he adopted Price's theory) was not a good economic reasoner. Still, even 150 years ago, there were bankers and Treasury officials, and it is possible that both they (and Price) were not so much bad theoretical reasoners as shrewd opportunists, that they were deliberately blind to the speciousness of an attractive defence of a desirable financial expedient. I have myself sometimes wondered whether, in the eighteenth century, an Assurance Society would have minded very much if a Life Table had erred on the pessimistic side.

Price did not enter on the population question with an unbiased mind. He was a keen politician and he believed that the policy of the government was bad for the country; he also believed that the wealth of a country was its people. Hence he believed that the population was declining and nothing shook that belief. Had he survived another ten years, until the first census, he would probably have disputed the accuracy of the returns.

Price began with the figures of houses in 1690, which he cited from Davenant (they were really due to King, who communicated them to Davenant), making the total 1,319,000. He then gave the figures of assessed, chargeable and cottages (cottages being houses too small to be taxed) as 678,915, 25,628 and 276,149, making a total of 980,692 in 1761. In 1777 they were 682,077, 19,396 and 251,261, a total of 952,734. On this basis he concluded that the population had declined by about one and a half millions and was actually less than five millions.

Howlett and Wales, Price's chief opponents, impugned every step in the reasoning. First, they pointed out that in the estimate for 1690 there was almost certainly a confusion between families and houses. Then they argued that many householders evaded duty (for instance by the simple plan of blocking up windows (the prayer 'Lighten our darkness, we beseech thee, Oh Pitt' is still remembered) and showed by direct enumeration in certain parishes that the returns were inaccurate. Finally, they gave reason to think that Price's multiplier was too small. On each of these points they were probably right. Indeed Price was obliged to admit the validity of some of their criticisms. But he declined to budge; sometimes he took *ad captandum* advantage of arithmetical slips by his adversaries, sometimes he declined to admit that their samples were representative, sometimes he tried to ignore the effect of corrections which he was forced to make.

These were the principal arguments. Both parties used the data of burials

and baptisms as subsidiary arguments. Price seems only to have used the London Bills, which rather let him down; because although they seemed to help for some part of the century, he admits that by 1773 London was increasing and, very characteristically, uses this as in his favour: 'But it appears that, in truth, this is an event more to be dreaded than desired. The more London increases, the more the rest of the country must be deserted.' Price's adversaries went farther afield and counted burials and christenings in 162 parishes in all parts of England for two quinquennia, one beginning in 1758, the other in 1773. Baptisms increased from 47,638 to 59,567, burials from 49,553 to 53,030.

But neither party put much weight upon what we should now consider primary evidence; rightly, because of its incompleteness.

But these data were not wholly neglected by medical writers as we shall see in later sections. One may fairly say on the evidence here summarized that the eighteenth-century political arithmeticians of England made no advance whatever upon the position reached by Graunt, Petty and King. They were second-rate imitators of men of genius.

(To be concluded)

MOMENTS OF THE DISTRIBUTIONS OF POWERS AND PRODUCTS OF NORMAL VARIATES

By J. B. S. HALDANE, F.R.S.

CONTENTS

		PAGE
1.	Introductory	226
2.	Distribution of the cube of a normal variate	228
3.	Distribution of any power of a normal variate	229
4.	Distribution of the product of three independent normal variates	230
5.	Distribution of the product of n independent normal variates	231
6.	Distribution of the product of two correlated normal variates	232
7.	Distribution of the product of three correlated normal variates	235
8.	The product of n correlated normal variates	238
9.	The Galton-Macalister distribution	239
10.	Biological applications	240
11.	Summary	241
	References	242

1. Introductory

When numerous organisms or organs are weighed, the distribution of the weights is often positively skew. On the other hand, the distributions of linear measurements is often very close to normal. The question then arises, given that the volume of an organ is proportional to the product of three mutually perpendicular measurements, each being normally distributed, what will be the distribution of the volumes? In general the three linear measurements will be correlated, and the problem might appear hopelessly complex. However, it will be shown later that, provided certain conditions are fulfilled, the distributions all lie very close together.

The problem can obviously be generalized to cover the case of the product of any number of normal variates. The most interesting cases are those of two, three, or an infinite number of such variates. Further, two special cases are comparatively simple. When the coefficients of correlation are all equal to unity and the coefficients of variation equal we are concerned with the distribution of a power of the normal variate. When the correlations all vanish, we are concerned with that of a product of several uncorrelated normal variates.

It is not, of course, suggested that all skew variation of weights is to be explained on these lines. For example the Galton-Macalister distribution, in which the logarithm of the variate is normally distributed, can be thought of as arising in at least two different ways. The weight may be the product of a large number of normal variates; or for constant cell size, the number of cell generations

may be distributed in a certain manner about a mean in each organ, these numbers being normally distributed. The highly skew variation of human weights is probably to be explained by the fact that a rather small fraction of the human race lays down very large quantities of fat. Nevertheless, it will be shown that simple criteria will determine whether observed positive skewness and lepto-kurtosis, too large to be ascribed to sampling error, can be explained on the lines discussed above. And in any particular case it is worth while finding out whether this is so.

Any measure of asymmetry, such as $\gamma_1 = \sqrt{\beta_1}$, or of kurtosis, such as $\gamma_2 = \beta_2 - 3$, is a dimensionless number independent of the unit of measurement. Hence in a transformed normal distribution of the type here considered it must clearly be a function of the only dimensionless number derivable from the first two moments, namely, the coefficient of variation, c. In what follows we shall generally use m for the mean of the original normal distribution and km^2 for its variance, so that $k^{\frac{1}{2}}$ is the coefficient of variation. The usual notation is used for the mean, variance and other moments and cumulants of the derived distribution.

The distribution of the product of a pair of correlated normal variates has already been fully discussed by Craig (1936). If m_1 and m_2 are the mean values of X and Y, k_1 and k_2 their coefficients of variation, and ρ their coefficients of correlation, then Craig finds for the cumulant generating function of

$$\frac{XY}{m_1 m_2 \sqrt{(k_1 k_2)}},$$

$$\kappa(\theta) = \frac{\theta\left[\left(\frac{1}{k_1}+\frac{1}{k_2}-\frac{2\rho}{\sqrt{(k_1k_2)}}\right)\theta+\frac{2}{\sqrt{(k_1k_2)}}\right]}{2[1-(1+\rho)\theta][1+(1-\rho)\theta]} - \tfrac{1}{2}\log[1-(1+\rho)\theta] - \tfrac{1}{2}\log[1+(1-\rho)\theta].$$

If $k_1 = k_2 = k$, this becomes

$$\kappa(\theta) = \frac{\theta}{k[1-(1+\rho)\theta]} - \tfrac{1}{2}\log[1-(1+\rho)\theta] - \tfrac{1}{2}\log[1+(1-\rho)\theta].$$

The rth cumulant of $\dfrac{XY}{m_1 m_2 \sqrt{(k_1 k_2)}}$ is

$$\kappa_r = \tfrac{1}{4}r!\left[\{(1+\rho)^{n-1}-(\rho-1)^{n-1}\}\left(\frac{1}{k_1}+\frac{1}{k_2}-\frac{2\rho}{\sqrt{(k_1k_2)}}\right)\right.$$
$$\left.+\{(1+\rho)^n-(\rho-1)^n\}\frac{2}{\sqrt{(k_1k_2)}}\right]+\tfrac{1}{2}(r-1)![(1+\rho)^n+(\rho-1)^n].$$

If $k_1 = k_2 = k$, the rth cumulant of XY is

$$\kappa_r = (m_1 m_2)^{r}k^{r-1}[(1+\rho)^{r-1}r! + \tfrac{1}{2}k\{(1+\rho)^r+(\rho-1)^r\}(r-1)!].$$

Craig's discussion is mainly confined to the cases where k_1 and k_2 are large. In the cases of greatest biometrical interest they are small, and some points of interest arise, besides those dealt with by Craig.

2. Distribution of the cube of a normal variate

If x be a reduced normal variate, that is to say, a variate whose mean is zero and standard deviation unity, it is required to find the distribution of X^3, where $X = m(1+k^{\frac{1}{2}}x)$. We note that

$$\overline{x^{2r+1}} = 0, \text{ and } \overline{x^{2r}} = (2r-1)(2r-3)(2r-5)\ldots 3.1 = \frac{(2r)!}{2^r . r!}.$$

Hence

$$\overline{X^3} = m^3(1+3k^{\frac{1}{2}}\bar{x}+3k\overline{x^2}+k^{\frac{1}{2}}\overline{x^3}) = m^3(1+3k),$$

$$\overline{X^6} = m^6(1+6k^{\frac{1}{2}}\bar{x}+15k\overline{x^2}+20k^{\frac{1}{2}}\overline{x^3}+15k^2\overline{x^4}+6k^{\frac{1}{2}}\overline{x^5}+k^3\overline{x^6}) = m^6(1+15k+45k^2+15k^3),$$

so that the first four moments of X^3 about zero are

$$\mu_1' = m^3(1+3k),$$

$$\mu_2' = m^6(1+15k+45k^2+15k^3),$$

$$\mu_3' = m^9(1+36k+378k^2+1260k^3+945k^4),$$

$$\mu_4' = m^{12}(1+66k+1485k^2+13860k^3+51975k^4+62370k^5+10395k^6).$$

Hence the moments about the mean $m^3(1+3k)$, and the cumulants, are

$$\left.\begin{aligned} \kappa_1 &= \mu_1' = m^3(1+3k), \\ \kappa_2 &= \mu_2 = 3m^6k(3+12k+5k^2), \\ \kappa_3 &= \mu_3 = 54m^9k^2(3+16k+15k^2), \\ \mu_4 &= 27m^{12}k^2(9+240k+1326k^2+1920k^3+385k^4), \\ \kappa_4 &= 648m^{12}k^3(7+48k+75k^2+15k^3). \end{aligned}\right\} \quad (1)$$

It is to be noted that in these calculations, and in the majority, though not all, those of this paper, the expressions for the moments about the mean are simpler than those for the moments about zero. Successive moments about the mean are therefore best calculated from the former, so far as possible, rather than the latter. Thus the equation

$$\mu_4 = \mu_4' - \mu_1'(4\mu_3+6\mu_1'\mu_2+\mu_1'^3),$$

involves less algebra than the more usual

$$\mu_4 = \mu_4' - 4\mu_1'\mu_3' + 6\mu_1'^2\mu_2 - 3\mu_1'^4.$$

It follows that

$$\sigma = 3m^3k^{\frac{1}{2}}(1+4k+\tfrac{5}{3}k^2)^{\frac{1}{2}},$$

$$c = \frac{(9k+36k^2+15k^3)^{\frac{1}{2}}}{1+3k},$$

$$\gamma_1 = \frac{+6\sqrt{(3k)}\,(3+16k+15k^2)}{(3+12k+5k^2)^{\frac{3}{2}}},$$

$$\gamma_2 = \frac{+72k(7+48k+75k^2+15k^3)}{(3+12k+5k^2)^2}.$$

Hence the distribution is positively skew, and leptokurtic. If c is small, we have, approximately,

$$\left.\begin{aligned}\gamma_1 &= 2c+\frac{2c^3}{27}+\ldots,\\ \gamma_2 &= \frac{56c^2}{9}+\frac{16c^4}{27}+\ldots.\end{aligned}\right\} \tag{2}$$

Since in most practical cases $c^2<0{\cdot}01$, the first term will have an error of less than 0·1 %, and no attention need be paid to the later terms.

3. Distribution of any power of a normal variate

It is required to find the moments of $m^n(1+k^{\frac{1}{2}}x)^n$. The mean value of its rth power is clearly

$$\mu'_r = m^{nr}\left[1+\frac{(nr)!\,k}{2(nr-2)!}+\frac{(nr)!\,k^2}{8(nr-4)!}+\ldots-\frac{(nr)!\,k^s}{2^s\,.\,s!(nr-2s)!}+\ldots\right].$$

This converges for all positive values of n, since k is small, though it only terminates for positive integral values of nr. After somewhat tedious algebra, we find for the mean and other moments and cumulants:

$$\left.\begin{aligned}\kappa_1 = \mu'_1 &= m^n[1+\tfrac{1}{2}n(n-1)\,k+\ldots],\\ \kappa_2 = \mu_2 &= m^{2n}n^2k[1+\tfrac{1}{2}(n-1)\,(3n-5)\,k+\ldots],\\ \kappa_3 = \mu_3 &= 3m^{3n}n^3(n-1)\,k^2[1+\tfrac{1}{6}(17n^2-55n+44)\,k+\ldots],\\ \mu_4 &= 3m^{4n}n^4k^2[1+\tfrac{5}{3}(n-1)\,(5n-7)\,k+\ldots],\\ \kappa_4 &= 4m^{4n}n^4(n-1)\,(4n-5)\,k^3[1+0(k)].\end{aligned}\right\} \tag{3}$$

Hence

$$\begin{aligned}c &= nk^{\frac{1}{2}}[1+\tfrac{1}{4}(n-1)\,(3n-5)\,k+\ldots],\\ \gamma_1 &= 3(n-1)\,k^{\frac{1}{2}}[1+\tfrac{1}{12}(7n^2-38n+43)\,k+\ldots],\\ \gamma_2 &= 4(n-1)\,(4n-5)\,k[1+0(k)].\end{aligned}$$

And when c is small

$$\left.\begin{aligned}\gamma_1 &= \frac{3(n-1)}{n}\,c[1+0(c^2)],\\ \gamma_2 &= \frac{4(n-1)\,(4n-5)\,c^2}{n^2}\,[1+0(c^2)].\end{aligned}\right\} \tag{4}$$

When $n=2$, the expressions become very simple,

$$\left.\begin{aligned}\kappa_1 &= m^2(1+k),\\ \kappa_2 &= 2m^4k(2+k),\\ \kappa_3 &= 8m^6k^2(3+k),\\ &\ldots\ldots\ldots\ldots\ldots\ldots\\ \kappa_r &= (r-1)!\,m^{2r}(2k)^{r-1}(r+k).\end{aligned}\right\} \tag{5}$$

This is in accordance with Craig's formula, putting $k_1=k_2=k$, $\rho=1$.

4. Distribution of the product of three independent normal variates

Let x, y, z, be independent reduced normal variates, so that $\overline{x^2} = \overline{y^2} = \overline{z^2} = 1$, $\overline{x^4} = \overline{y^4} = \overline{z^4} = 3$, etc. Let $X = m_1(1 + k_1^1 x)$, $Y = m_2(1 + k_2^1 y)$, $Z = m_3(1 + k_3^1 z)$. Here k_1^1, k_2^1, k_3^1 are coefficients of variation of linear measurements, so that k_1, k_2, k_3, are commonly about 0·001. It is required to find the moments of the distribution of XYZ. In the expansion of a power of XYZ we need only consider even powers of x, y and z. For example,

$$\begin{aligned}\overline{X^2Y^2Z^2} &= m_1^2 m_2^2 m_3^2 \overline{(1+k_1^1 x)^2 (1+k_2^1 y)^2 (1+k_3^1 z)^2} \\ &= m_1^2 m_2^2 m_3^2 (1+k_1\overline{x^2})(1+k_2\overline{y^2})(1+k_3\overline{z^2}) \\ &= m_1^2 m_2^2 m_3^2 (1+k_1)(1+k_2)(1+k_3).\end{aligned}$$

So if $m_1 m_2 m_3 = V$ (a volume), the moments of XYZ about zero are

$$\begin{aligned}\mu_1' &= V, \\ \mu_2' &= V^2(1+k_1)(1+k_2)(1+k_3), \\ \mu_3' &= V^3(1+3k_1)(1+3k_2)(1+3k_3), \\ \mu_4' &= V^4(1+6k_1+3k_1^2)(1+6k_2+3k_2^2)(1+6k_3+3k_3^2).\end{aligned}$$

So
$$\begin{aligned}\kappa_2 = \mu_2 &= V^2(\Sigma k_1 + \Sigma k_1 k_2 + k_1 k_2 k_3), \\ \kappa_3 = \mu_3 &= 6V^3(\Sigma k_1 k_2 + 4k_1 k_2 k_3), \\ \mu_4 &= 3V^4(\Sigma k_1^2 + 2\Sigma k_1 k_2 + 6\Sigma k_1^2 k_2 + 38 k_1 k_2 k_3 + 3\Sigma k_1^2 k_2^2 \\ &\qquad + 36\Sigma k_1^2 k_2 k_3 + 18\Sigma k_1^2 k_2 k_3 + 9 k_1^2 k_2^2 k_3^2, \\ \kappa_4 &= 6V^4[2\Sigma k_1^2 k_2 + \Sigma k_1^2 k_2^2 + 4k_1 k_2 k_3(4 + 4\Sigma k_1 + 2\Sigma k_1 k_2 + k_1 k_2 k_3)].\end{aligned}$$

In practice we can neglect all terms except the leading one, and write

$$\left.\begin{aligned}\kappa_1 &= V, \\ \kappa_2 &= V^2(k_1 + k_2 + k_3), \\ \kappa_3 &= 6V^3(k_2 k_3 + k_3 k_1 + k_1 k_2), \\ \kappa_4 &= 12V^4(k_1^2 k_2 + k_2^2 k_3 + k_3^2 k_1 + k_1 k_2^2 + k_2 k_3^2 + k_3 k_1^2 + 8 k_1 k_2 k_3).\end{aligned}\right\} \tag{6}$$

If $k_1 = k_2 = k_3 = k$, we find

$$\left.\begin{aligned}\kappa_1 &= V, \\ \kappa_2 &= V^2 k(3 + 3k + k^2), \\ \kappa_3 &= 6V^3 k^2(3 + 4k), \\ \kappa_4 &= 6V^4 k^3(28 + 51k + 24k^2 + 4k^3),\end{aligned}\right\} \tag{7}$$

which may be compared with equations (1). Thus

$$\left.\begin{aligned}c &= \sqrt{(3k + 3k^2 + k^3)}, \\ \gamma_1 &= 2c - \frac{4c^3}{9} + \ldots, \\ \gamma_2 &= \frac{56c^2}{9} - \frac{22c^4}{9} + \ldots.\end{aligned}\right\} \tag{8}$$

It will be noticed that the leading terms of equations (2) and (8) are identical. Hence the two distributions will, in practice, be quite indistinguishable, even if tens of thousands of individuals are weighed. It follows that the cube root of the product of three independent variates is almost normally distributed. Further, the relations of γ_1 and γ_2 to c are little altered when k_1, k_2 and k_3 are different, provided that they are of the same order of magnitude. Thus if $k_1 = 0{\cdot}001$, $k_2 = 0{\cdot}002$, $k_3 = 0{\cdot}003$, we find $\gamma_1 = \frac{11}{3}c$ instead of $2c$, and $\gamma_2 = \frac{16c^2}{3} = 5{\cdot}\dot{3}c^2$ instead of $6{\cdot}\dot{2}c^2$. Only a very large sample indeed would reveal departures of this order from the simpler expressions of equations (2) and (8).

5. Distribution of the product of n independent normal variates

Let $x_1, x_2, x_3, \ldots, x_r, \ldots, x_n$, be n independent reduced normal variates. The general case is of course somewhat complicated, so we shall only investigate the special case where all the coefficients of variation are equal.

Then if $X_r = m_r(1 + k^{\frac{1}{2}}x_r)$, and $M = \Pi m_r$, the moments about zero of the distribution of the product ΠX_r are

$$\begin{aligned}
\mu'_1 &= M,\\
\mu'_2 &= M^2(1+k)^n,\\
\mu'_3 &= M^3(1+3k)^n,\\
\mu'_4 &= M^4(1+6k+3k^2)^n,\\
\mu'_5 &= M^5(1+10k+15k^2)^n,\\
\mu'_6 &= M^6(1+15k+45k^2+15k^3)^n.
\end{aligned}$$

Hence

$$\mu_2 = M^2[(1+k)^n - 1],$$

$$\mu_3 = M^3[(1+3k)^n - 3(1+k)^n + 2],$$

$$\mu_4 = M^4[(1+6k+3k^2)^n - 4(1+3k)^n + 6(1+k)^n - 3],$$

$$\mu_5 = M^5[(1+10k+15k^2)^n - 5(1+6k+3k^2)^n + 10(1+3k)^n - 10(1+k)^n - 4],$$

$$\mu_6 = M^6[(1+15k+45k^2+15k^3)^n - 6(1+10k+15k^2)^n + 15(1+6k+3k^2)^n - 20(1+3k)^n + 15(1+k)^n - 5],$$

$$\kappa_4 = M^4[(1+6k+3k^2)^n - 3(1+2k+k^2)^n - 4(1+3k)^n + 12(1+k)^n - 6],$$

$$\kappa_5 = M^5\{(1+10k+15k^2)^n + 5[-(1+6k+3k^2)^n - 2(1+4k+3k^2)^n + 6(1+2k+k^2)^n + 4(1+3k)^n - 12(1+k)^n] + 24\},$$

$$\begin{aligned}
\kappa_6 = M^6\{&(1+15k+45k^2+15k^3)^n + 15[2(1+3k+3k^2+k^3)^n\\
&- (1+7k+9k^2+3k^3)^n + 2(1+6k+3k^2)^n + 8(1+4k+3k^2)^n\\
&- 18(1+2k+k^2)^n - 8(1+3k)^n + 24(1+k)^n - 8]\\
&- 6(1+10k+15k^2)^n - 10(1+6k+9k^2)^n\}.
\end{aligned}$$

The higher cumulants can be evaluated as follows:

$$\kappa_3 = 3M^3 \sum_{r=1}^{\infty} \binom{n}{r} (3^{r-1}-1)\, k^r,$$

$$\kappa_4 = 3M^4 \sum_{r=1}^{\infty} \binom{n}{r} (3^{r-1}-1)\,[(2+k)^r - 4]\, k^r, \text{ etc.}$$

Thus we find

$$\left.\begin{aligned}
&\kappa_1 = M,\\
&\kappa_2 = M^2 nk[1 + \tfrac{1}{2}(n-1)\,k + \ldots],\\
&\kappa_3 = M^3 n(n-1)\,k^2[3 + 4(n-2)\,k + \ldots],\\
&\kappa_4 = M^4 n(n-1)\,k^3[4(4n-5) + 3(13n^2 - 49n + 47)\,k + \ldots],\\
&\kappa_5 = 5M^5 n(n-1)\,k^4[(5n-6)(5n-7) + 2(n-2)(143n-345)(k+\ldots],\\
&\kappa_6 = 3M^6 n(n-1)\,k^5[432n^3 - 1853n^2 + 2917n - 1758 + 0(k)].
\end{aligned}\right\} \quad (9)$$

These may be compared with equations (3).

$$c = \sqrt{(nk)}\,[1 + \tfrac{1}{4}(n-1)\,k + \ldots],$$

and when c is small

$$\gamma_1 = \frac{3(n-1)\,c}{n}\,[1 + 0(c^2)],$$

$$\gamma_2 = \frac{4(n-1)(4n-5)\,c^2}{n^2}\,[1 + 0(c^2)],$$

as in equations (4).

When $n = 2$ we have the simple forms

$$\kappa_1 = M,$$

$$\kappa_2 = M^2 k(2+k),$$

$$\kappa_3 = 6M^3 k^2,$$

$$\kappa_4 = 6M^4 k^3 (4+k),$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\kappa_{2r} = (2r-1)!\, M^{2r} k^{2r-1}(2r+k),$$

$$\kappa_{2r+1} = (2r+1)!\, M^{2r+1} k^{2r}.$$

This is in accordance with Craig's formula, putting $k_1 = k_2 = k$, $\rho = 0$.

6. Distribution of the product of two correlated normal variates

In general the different linear dimensions of an organ or organism are positively correlated. Organic correlations may reach very high values, such as 0·9, and presumably even higher values would be found for two approximately equal diameters of an approximate solid of revolution, such as an apple or an egg, where ρ may be taken as unity. On the other hand, quite low values are found. Thus the

length and breadth of a homogeneous group of like-sexed adult human skulls generally show a correlation of about 0·3 or 0·4. Hence the area of an organ will commonly be proportional to the product of two correlated variables, the volume to the product of three.

Let x and y be two reduced normal variates, with correlation ρ. The cumulant-generating function for their joint distribution is $\frac{1}{2}(t^2+2\rho tu+u^2)$. That is to say if $r+s$ is odd, then $\overline{x^r y^s}=0$. If $r+s$ is even, then $\overline{x^r y^s}$ is the coefficient of $\frac{t^r u^s}{r!\,s!}$ in $\exp\frac{1}{2}(t^2+2\rho tu+u^2)$. That is to say, $\overline{x^r y^s}$ is the coefficient of $t^r u^s$ in

$$(t^2+2\rho tu+u^2)^{\frac{1}{2}(r+s)},$$

multiplied by

$$\frac{r!\,s!}{2^{\frac{1}{2}(r+s)}[\frac{1}{2}(r+s)]!}.$$

It follows that

$$\overline{x^2}=1, \quad \overline{xy}=\rho,$$

$$\overline{x^4}=3, \quad \overline{x^3y}=3\rho, \quad \overline{x^2y^2}=1+2\rho^2,$$

$$\overline{x^6}=15, \quad \overline{x^5y}=15\rho, \quad \overline{x^4y^2}=3(1+4\rho^2), \quad \overline{x^3y^3}=3\rho(3+2\rho^2),$$

$$\overline{x^8}=105, \quad \overline{x^7y}=105\rho, \quad \overline{x^6y^2}=15(1+6\rho^2), \quad \overline{x^5y^3}=15\rho(3+4\rho^2),$$

$$\overline{x^4y^4}=3(3+24\rho^2+8\rho^4),$$

$$\overline{x^8y^2}=105(1+8\rho^2), \quad \overline{x^7y^3}=315\rho(1+2\rho^2), \quad \overline{x^6y^4}=45(1+12\rho^2+8\rho^4),$$

$$\overline{x^8y^4}=315(1+16\rho^2+16\rho^4).$$

Of course $\overline{x^s y^r}=\overline{x^r y^s}$.

Thus the moments of xy about its mean ρ are

$$\mu_2=1+\rho^2, \quad \mu_3=2\rho(3+\rho^2), \quad \mu_4=3(3+14\rho^2+3\rho^4).$$

Hence $\kappa_4=6(1+6\rho^2+\rho^4)$. Hence the distribution is leptokurtic, and asymmetrical unless ρ vanishes.

Now consider two correlated normal variates

$$X=m_1(1+ax), \quad \text{and} \quad Y=m_2(1+by),$$

where a and b are coefficients of variation. Let $m_1m_2=A$. Then the moments of XY about zero are

$$\mu_1'=A(1+\rho ab),$$

$$\mu_2'=A^2[1+a^2+4\rho ab+b^2+(1+2\rho^2)a^2b^2],$$

$$\mu_3'=A^3[1+3(a^2+3\rho ab+b^2)+9ab\{\rho a^2+(1+2\rho^2)ab+\rho b^2\}+3\rho(3+2\rho^2)a^3b^3],$$

$$\mu_4'=A^4[1+2(3a^2+8\rho ab+3b^2)+3\rho a^4+16\rho a^3b+12(1+2\rho^2)a^2b^2+16\rho ab^3+3\rho b^4$$
$$+6a^2b^2\{3(1+4\rho^2)a^2+8\rho(3+2\rho^2)ab+3(1+4\rho^2)b^2\}+3(3+24\rho^2+8\rho^4)a^4b^4].$$

Hence

$$\kappa_2 = \mu_2 = A^2[a^2 + 2\rho ab + b^2 + (1+\rho^2)a^2b^2],$$

$$\kappa_3 = \mu_3 = 2A^3ab[3\{\rho a^2 + (1+\rho^2)ab + \rho b^2\} + \rho(3+\rho^2)a^2b^2],$$

$$\mu_4 = 3A^3[(a^2+2\rho ab+b^2)^2 + 2a^2b^2\{(3+7\rho^2)a^2 + 2\rho(7+3\rho^2)ab + (3+7\rho^2)b^2\} + (3+14\rho^2+3\rho^4)a^4b^4],$$

$$\kappa_4 = 6A^3a^2b^2[2\{(1+3\rho^2)a^2 + 2\rho(3+\rho^2)ab + (1+3\rho^2)b^2\} + (1+6\rho^2+\rho^4)a^2b^2], \quad (10)$$

in accordance with Craig's formula.

The most interesting case arises when $a = b = k^{\frac{1}{2}}$. This case is important because in practice the coefficients of variation of different linear dimensions of the same organ are often nearly equal. Thus those of linear skull measurements in like-sexed adults in a racially homogeneous population are about 0·03, so that k is about 0·001. The moments and cumulants of XY are then

$$\kappa_1 = A(1+\rho k),$$

$$\kappa_2 = \mu_2 = A^2k[2(1+\rho) + (1+\rho^2)k],$$

$$\kappa_3 = \mu_3 = 2A^3k^2[3(1+\rho)^2 + (3+\rho^2)k],$$

$$\mu_4 = 3A^4k^2[4(1+\rho)^2 + 4(3+7\rho+7\rho^2+3\rho^3)k + (3+14\rho^2+3\rho^4)k^2],$$

$$\kappa_4 = 6A^4k^3[4(1+\rho)^3 + (1+6\rho^2+\rho^4)k]. \quad (11)$$

Hence

$$c = \sqrt{[2(1+\rho)k]}\left[1 + \frac{1-4\rho-3\rho^2}{4(1+\rho)}k + \ldots,\right]$$

while

$$\gamma_1 = \tfrac{3}{2}c[1+0(c^2)],$$

$$\gamma_2 = 3c^2[1+0(c^2)],$$

exactly as found when the two variates are in a constant ratio or quite independent. When however $a \neq b$, this is no longer the case, for it is clear that the distribution becomes normal when a or b vanishes. If $\dfrac{a^2+b^2}{2ab} = p$, so that $p \geqslant 1$, we have, for equations (10),

$$\gamma_1 = \frac{1+2p\rho+\rho^2}{(p+\rho)^2}\frac{3c}{2}, \quad \gamma_2 = \frac{1+3p\rho-3\rho^2+p\rho^3}{(p+\rho)^3}3c^2,$$

both approximately. However, a and b may differ considerably without any very great effect of γ_1/c or γ_2/c. Thus if $a = 2b$, so that $p = \frac{5}{4}$,

$$\gamma_1 = \frac{8(2+5\rho+2\rho^2)}{(5+4\rho)^2}\frac{3c}{2}.$$

That is to say, γ_1 is 64 % of $3c/2$ if $\rho = 0$, and 86 % if $\rho = \frac{1}{2}$, whilst γ_2 is 51 % of $3c$ if $\rho = 0$, and 71 % if $\rho = \frac{1}{2}$.

So far we have assumed that ρ is not negative. If $\rho = -1$, $XY = 1 - kx^2$, so the mean is $A(1-k)$, and the other cumulants are given by

$$\kappa_r = (-)^r\, 2^{r-1}(r-1)!\, A^r k^r.$$

If $\rho = -1 + 2\sqrt{\frac{k}{3} + \frac{k}{2}} + \ldots$, κ_3 in equations (11) vanishes, though the distribution does not become quite symmetrical. However, negative values of ρ are of no biological interest.

7. Distribution of the product of three correlated normal variates

Let x, y, z, be three reduced normal variates as before, their correlations being $\rho_{yz} = \kappa$, $\rho_{zx} = \lambda$, $\rho_{xy} = \mu$. Thus the cumulant-generating function is

$$\tfrac{1}{2}(t^2 + u^2 + v^2 + 2\kappa uv + 2\lambda vt + 2\mu tu).$$

Odd moments vanish, and the even moment $\overline{x^p y^q z^r}$ is the coefficient of $t^p u^q v^r$ in

$$(t^2 + u^2 + v^2 + 2\kappa uv + 2\lambda vt + 2\mu tu)^{\frac{1}{2}(p+q+r)},$$

multiplied by $$\frac{p!\,q!\,r!}{2^{\frac{1}{2}(p+q+r)}\,[\frac{1}{2}(p+q+r)]!}.$$

The required moments of products of two variates, such as

$$\overline{xy} = \mu, \quad \overline{x^4y^2} = 3(1 + 4\mu^2),$$

have already been given. The required moments of products of all three are:

$$\overline{x^2yz} = \kappa + 2\lambda\mu,$$

$$\overline{x^4yz} = 3(\kappa + 4\lambda\mu), \quad \overline{x^3y^2z} = 3(\lambda + 2\kappa\mu + 2\kappa\mu^2), \quad \overline{x^2y^2z^2} = 1 + 2\Sigma\kappa^2 + 8\kappa\lambda\mu,$$

$$\overline{x^4y^3z} = 3(3\kappa + 12\lambda\mu + 12\kappa\mu^2 + 8\lambda\mu^3),$$

$$\overline{x^4y^2z^2} = 3(1 + 2\kappa^2 + 4\lambda^2 + 4\mu^2 + 16\kappa\lambda\mu + 8\lambda^2\mu^2),$$

$$\overline{x^3y^3z^2} = 3(3\mu + 6\kappa\lambda + 2\mu^3 + 6\kappa^2\mu + 6\lambda^2\mu + 12\kappa\lambda\mu^2),$$

$$\overline{x^4y^4z^2} = 3(3\kappa + 12\kappa^2 + 12\lambda^2 + 24\mu^2 + 96\kappa\lambda\mu + 48\kappa^2\mu^2 + 48\lambda^2\mu^2 + 8\mu^4 + 64\kappa\lambda\mu^3),$$

$$\overline{x^4y^3z^3} = 9(3\kappa + 12\lambda\mu + 2\kappa^3 + 12\kappa\lambda^2 + 12\kappa\mu^2 + 24\kappa^2\lambda\mu + 8\lambda^3\mu + 8\lambda^3\mu + 8\lambda\mu^3 + 24\kappa\lambda^2\mu^2),$$

$$\overline{x^4y^4z^4} = 9[3 + 24\Sigma\kappa^2 + 8\Sigma\kappa^4 + 96\Sigma\lambda^2\mu^2 + 64\kappa\lambda\mu(3 + 2\Sigma\kappa^2 + 3\kappa\lambda\mu)].$$

Other moments (except those such as $\overline{xy^2z} = \lambda + 2\kappa\mu$, which are derivable by transposition) are not needed for our purposes. It follows that xyz has a symmetrical and leptokurtic distribution, with mean zero, and

$$\kappa_2 = \mu_2 = 1 + 2\Sigma\kappa^2 + 8\kappa\lambda\mu,$$

$$\kappa_4 = 12[2 + 17\Sigma\kappa^2 + 5\Sigma\kappa^4 + 70\Sigma\lambda^2\mu^2 + 4\kappa\lambda\mu(35 + 22\Sigma\kappa^2 + 32\kappa\lambda\mu)].$$

If $X = m_1(1 + ax)$, $Y = m_2(1 + by)$, $Z = m_3(1 + cz)$ the expressions for the higher moments are very complicated, though it can easily be shown that the mean is $m_1 m_2 m_3(1 + \kappa bc + \lambda ca + \mu ab)$, and the variance

$$m_1^2 m_2^2 m_3^2[\Sigma a^2 + 2\Sigma\kappa bc + \Sigma(1 + \kappa^2)\,b^2c^2 + 2\Sigma(2\kappa + 3\lambda\mu)\,a^2bc + (1 + 2\Sigma\kappa^2 + 8\kappa\lambda\mu)\,a^2b^2c^2].$$

We shall only give further consideration to the case where all the coefficients of variation are equal. And we shall confine ourselves to two special cases of this. In the one $\kappa = \lambda = \mu = \rho$, that is to say the variates are equally, and therefore of course positively, correlated. This is not very far from the case with the human skull. In the other, $\kappa = 1$, whilst $\lambda = \mu = \rho$. This is appropriate to an approximate solid of revolution, such as many eggs and fruits, or to a regular prism such as some sponge spicules.

In the first case we have for the moments of products of three variables:

$$\overline{x^2yz} = \rho(1+2\rho),$$

$$\overline{x^4yz} = 3\rho(1+4\rho), \quad \overline{x^3y^2z} = 3\rho(1+2\rho+2\rho^2), \quad \overline{x^2y^2z^2} = 1+6\rho^2+8\rho^3,$$

$$\overline{x^4y^3z} = 3\rho(3+12\rho+12\rho^2+8\rho^3), \quad \overline{x^4y^2z^2} = 3(1+10\rho^2+16\rho^3+8\rho^4),$$

$$\overline{x^3y^3z^3} = 3\rho(3+6\rho+14\rho^2+12\rho^3),$$

$$\overline{x^4y^4z^2} = 3(3+48\rho^2+96\rho^3+104\rho^4+64\rho^5),$$

$$\overline{x^4y^3z^3} = 9\rho(3+12\rho+26\rho^2+40\rho^3+24\rho^4),$$

$$\overline{x^4y^4z^4} = 27(1+24\rho^2+64\rho^3+104\rho^4+128\rho^5+64\rho^6).$$

So if $X = m_1(1+k^{\frac{1}{2}}x)$, $Y = m_2(1+k^{\frac{1}{2}}y)$, $Z = m_3(1+k^{\frac{1}{2}}z)$, and $m_1m_2m_3 = V$, the volume which is the product of the means, then the moments of XYZ about zero are

$$\mu'_1 = V(1+3\rho k),$$

$$\mu'_2 = V^2[1+3(1+4\rho)k+3(1+4\rho+10\rho^2)k+(1+6\rho^2+8\rho^3)k^2],$$

$$\mu'_3 = V^3[1+9(1+3\rho)k+27(1+5\rho+8\rho^2)k^2+9(3+21\rho+54\rho^2 + 62\rho^3)k^3+27\rho(3+6\rho+14\rho^2+12\rho^3)k^4],$$

$$\mu'_4 = V^4[1+6(3+8\rho)k+9(13+64\rho+88\rho^2)k^2+36(9+64\rho+160\rho^2 + 152\rho^3)k^3+27(13+128\rho+448\rho^2+768\rho^3+568\rho^4)k^4 + 54(3+24\rho+144\rho^2+304\rho^3+424\rho^4+256\rho^5)k^5 + 27(1+24\rho^2+64\rho^3+104\rho^4+128\rho^5+64\rho^6)k^6].$$

Hence

$$\kappa_1 = \mu'_1 = V(1+3\rho k),$$

$$\kappa_2 = \mu_2 = V^2k[3(1+2\rho)+3(1+4\rho+7\rho^2)k+(1+6\rho^2+8\rho^3)k^2],$$

$$\kappa_3 = \mu_3 = 6V^3k^2[3(1+2\rho)^2+(41+27\rho+60\rho^2+53\rho^3)k + 3\rho(4+9\rho+8\rho^2+14\rho^3)k^2],$$

$$\mu_4 = 3V^4k^2[9(1+2\rho)^2+2(37+222\rho+471\rho^2+350\rho^3)k+3(39+316\rho + 1038\rho^2+1584\rho^3+1001\rho^4)k^2+8(3+24\rho+127\rho^2+268\rho^3 + 346\rho^4+192\rho^5)k^3+9(1+24\rho^2+64\rho^3+104\rho^4+128\rho^5+64\rho^6)k^4],$$

$$\kappa_4 = 6V^4k^3[28(1+2\rho)^3+3(17+144\rho+468\rho^2+688\rho^3+411\rho^4)k + 12(2+17\rho+92\rho^2+193\rho^3+241\rho^4+130\rho^5)k^2 + 2(2+51\rho^2+140\rho^3+225\rho^4+264\rho^5+128\rho^6)k^3]. \quad (12)$$

These expressions reduce to equations (1) if $\rho = 1$, and (7) if $\rho = 0$. When k is small, $c^2 = 3(1+2\rho)k$,

$$\gamma_1 = 2c[1+0(c^2)],$$

$$\gamma_2 = \frac{56c^2}{9}[1+0(c^2)].$$

Thus the relation of γ_1 and γ_2 to c is almost independent of the coefficient of correlation.

We next consider the case when $z = y$, so that $\kappa = 1$, whilst $\lambda = \mu = \rho$. This is most simply solved by finding the moments of $(1+k^{\frac{1}{2}}x)(1+k^{\frac{1}{2}}y)^2$. Thus if $X = m_1(1+k^{\frac{1}{2}}x)$, and $Y = m_2(1+k^{\frac{1}{2}}y)$, the moments of XY^2 about zero are

$$\mu_1' = V[1+(1+2\rho)k],$$

$$\mu_2' = V^2[1+(7+8\rho)k+3(3+8\rho+4\rho^2)k^2+3(1+4\rho^2)k^3],$$

$$\mu_3' = V^3[1+18(1+\rho)k+18(5+11\rho+5\rho^2)k^2+30(5+15\rho+18\rho^2+4\rho^3)k^3 + 45(1+6\rho+6\rho^2+8\rho^3)k^4],$$

$$\mu_4' = V^4[1+2(17+16\rho)k+3(127+256\rho+112\rho^2)k^2+84(21+64\rho+64\rho^2 + 16\rho^3)k^3+105(31+128\rho+192\rho^2+128\rho^3+16\rho^4)k^4 + 630(3+16\rho+32\rho^2+32\rho^3+16\rho^4)k^5+315(1+16\rho^2+16\rho^4)k^6].$$

Hence

$$\kappa_1 = \mu_1' = V[1+(1+2\rho)k],$$

$$\kappa_2 = \mu_2 = V^2k[5+4\rho+4(2+5\rho+2\rho^2)k+3(1+4\rho^2)k^2],$$

$$\kappa_3 = \mu_3 = 2V^3k^2[3(8+14\rho+5\rho^2)+2(29+84\rho+87\rho^2+16\rho^3)k + 9(2+14\rho+13\rho^2+16\rho^3)k^2],$$

$$\mu_4 = 3V^4k^2[(5+4\rho)^2+8(40+113\rho+95\rho^2+22\rho^3)k+2(427 + 1628\rho+2376\rho^2+1376\rho^2+160\rho^3+160\rho^4)k^2 + 24(24+121\rho+237\rho^2+234\rho^3+104\rho^4)k^3+105(1+16\rho^2+16\rho^4)k^4],$$

$$\kappa_4 = 24V^4k^3[30+80\rho+65\rho^2+14\rho^3+(95+364\rho+513\rho^2+292\rho^3 + 32\rho^4)k+3(22+116\rho+227\rho^2+214\rho^3+96\rho^4)k^2 + 3(4+67\rho^2+64\rho^4)k^3]. \quad (13)$$

Hence, approximately,

$$c^2 = (5+4\rho)k,$$

$$\gamma_1 = \frac{6(2+\rho)(4+5\rho)c}{(5+4\rho)^2},$$

$$\gamma_2 = \frac{24(30+80\rho+65\rho^2+14\rho^3)c^2}{(5+4\rho)^3}.$$

Hence γ_1 varies between $\frac{48}{25}c$, or $1{\cdot}92c$ when $\rho = 0$, and $2c$ when $\rho = 1$. This is its maximum, so for high values of ρ, γ_1/c is nearly constant. It vanishes when $\rho = -0{\cdot}8$. γ_2 increases from $\frac{144}{25}c^2$, or $5{\cdot}76c^2$, when $\rho = 0$, to $\frac{56}{9}c^2$, or $6{\cdot}2\dot{2}c^2$ when

$\rho = 1$. This again is its maximum, so γ_2/c^2 is nearly constant for large ρ. γ_2 does not vanish for any admissible values of ρ. In fact for positive values of ρ it would be impossible except with enormous samples, to distinguish this distribution from that of the cube of a normal variate. If $\rho = 1$, equations (13) reduce to equations (1). If $\rho = 0$, they give the cumulants of XY^2, where X and Y are uncorrelated normal variates with the same coefficient of variation, namely,

$$\kappa_1 = V(1+k), \qquad \kappa_2 = V^2k(5+8k+3k^2),$$

$$\kappa_3 = 4V^3k^2(12+29k+9k^2), \quad \kappa_4 = 24V^4k^3(30+95k+66k^2+12k^3),$$

which may easily be obtained independently.

8. The product of n correlated normal variates

If the variates are X_1, X_2, ..., X_r, ..., X_n, where $X_r = m_r(1+a_rx_r)$, and $P = \Pi m_r$, while x_r is a reduced normal variate, and ρ_{rs} is the coefficient of correlation of x_r and x_s, and hence of X_r and X_s, the general expression for the moments of $X_1, X_2, \ldots, X_n$ is complicated. It can however easily be seen that the mean is:

$$P[1+\Sigma\rho_{rs}a_ra_s+\Sigma(\rho_{rs}\rho_{tu}+\rho_{rt}\rho_{su}+\rho_{ru}\rho_{ts})a_ra_sa_ta_u+\ldots],$$

while

$$\mu_2 = P^2[\Sigma a_r^2+2\Sigma\rho_{rs}a_ra_s+\ldots].$$

If every $a_r = k^{\frac{1}{2}}$, and every $\rho_{rs} = \rho$, then $\overline{x_1^{z_1}x_2^{z_2}x_3^{z_3}\ldots}$ vanishes when $m = \Sigma\alpha_r$ is odd, and when m is even it is the coefficient of $t_1^{z_1}t_1^{z_2}\ldots$ in the expansion of $(\Sigma t_r^2+2\rho\Sigma t_rt_s)^m$, multiplied by $\alpha_1!\alpha_2!\alpha_3!/(2^mm!)$. Thus the moments about zero are:

$$\mu_1' = P\left[1+\tfrac{1}{2}n(n-1)\rho k+\tfrac{1}{8}n(n-1)(n-2)(n-3)\rho^2k^2+\ldots+\frac{n!\rho^rk^r}{2^rr!(n-2r)!}+\ldots\right],$$

$$\mu_2' = P^2[1+n\{1+2(n-1)\rho\}k+\tfrac{1}{2}n(n-1) \times\{1+4(n-2)\rho+2(2n^2-6n+5)\rho^2\}k^2+\ldots],$$

$$\mu_3' = P^3[1+3n\{1+\tfrac{3}{2}(n-1\rho\}k+\tfrac{9}{2}n(n-1) \times\{1+(3n-4)\rho+\tfrac{1}{4}(9n^2-21n+14)\rho^2\}k^2+\ldots],$$

So

$$\mu_2 = P^2nk[1+(n-1)\rho+\tfrac{1}{2}(n-1)\{1+4(n-2)\rho+(3n^2-9n+7)\rho^2\}k+\ldots],$$

$$\mu_3 = 3P^2n(n-1)[1+(n-1)\rho^2k^2[1+0(k)]. \tag{14}$$

Thus

$$c = [n\{1+(n-1)\rho\}k]^{\frac{1}{2}}[1+0(k)],$$

$$\gamma_1 = 3(n-1)\left[\frac{1+(n-1)\rho}{n}\right]^{\frac{1}{2}}[1+0(k)] \tag{15}$$

$$= \frac{3(n-1)}{n}c^2[1+0(c^2)],$$

and presumably γ_2 etc. are approximately the same functions of c as in the case of equations (9).

9. THE GALTON-MACALISTER DISTRIBUTION

Consider the distribution of the nth power of a normal variate when n is very large, and k very small, but $nk^{\frac{1}{2}}$ remains constant. All but a few terms of equations (3) will vanish, and if $nk^{\frac{1}{2}}$ is small, the distribution becomes identical with the distribution described by MacAlister (1879) whose first four moments have been given by Pearson (1905). Its cumulants can readily be found as follows: Given that $x = \log X$ is normally distributed, to find the distribution of X. Let m and s be the mean and standard deviation of x. Then $X = e^x$. Hence, since the moment-generating function of x is $M(t) = e^{mt+\frac{1}{2}st^2}$, the rth moment of X about zero is

$$\mu'_r = \overline{e^{rx}} = 1 + r\mu_1 + \frac{r^2\mu_2}{2!} + \ldots = M(r) = e^{mr+\frac{1}{2}sr^2},$$

where μ_1, μ_2, etc. are the moments of x about zero. Let $M = e^{m+\frac{1}{2}s}$, $l = e^s$. Then $\mu'_r = M^r l^{\frac{1}{2}r(r-1)}$. Hence the cumulants of X are

$$\begin{aligned}
\kappa_1 &= M,\\
\kappa_2 &= M^2(l-1),\\
\kappa_3 &= M^3(l-1)^2(l+2),\\
\kappa_4 &= M^4(l-1)^3(l^3+3l^2+6l+6),\\
\kappa_5 &= M^5(l-1)^4(l^6+4l^5+10l^4+20l^3+30l^2+36l+24),\\
\kappa_6 &= M^6(l-1)^5(l^{10}+5l^9+15l^8+35l^7+70l^6+120l^5+180l^4\\
&\qquad +240l^3+270l^2+240l+120),
\end{aligned} \tag{16}$$

or, if $l-1$ is a small quantity, q, approximating to s,

$$\begin{aligned}
\kappa_1 &= M,\\
\kappa_2 &= M^2q,\\
\kappa_3 &= M^3q^2(3+q),\\
\kappa_4 &= M^4q^3(16+15q+6q^2+q^3),\\
\kappa_5 &= M^5q^4(125+222q+205q^2+120q^3+45q^4+10q^5+q^6),\\
\kappa_6 &= M^6q^5(1296+3660q+5700q^2+5165q^3+4945q^4+2997q^5\\
&\qquad +1365q^6+455q^7+105q^8+15q^9+q^{10}).
\end{aligned}$$

Hence

$$\left.\begin{aligned}
c^2 &= q,\\
\gamma_1 &= 3c+c^3,\\
\gamma_2 &= 16c^2+15c^4+\ldots,\\
\gamma_3 &= 125c^3+222c^5+\ldots,\\
\gamma_4 &= 1296c^4+3660c^6+\ldots.
\end{aligned}\right\} \tag{17}$$

The first terms represent the limiting values of equations (3) and (9). The leading term of γ_r up to γ_4 at least, is $(r+2)^r$. Thus for a given coefficient of variation, γ_1 is 50 % greater than in the case of the cubed normal variate, γ_2 is 157 % greater.

10. Biological applications

Wilson and Hilferty (1931) showed that the cube root of χ^2 is almost normally distributed when n exceeds 2. Haldane (1938) gave the value of the cumulants in this case, and showed that, for large values of n, $(13\chi^2-n)^{\frac{1}{3}}$ is even more nearly normally distributed. It is interesting to compare the cumulants of the cubed-normal distribution [equations (2)] with those of the χ^2 distribution. The latter are $\kappa_1 = n$, $\kappa_2 = 2n$, $\kappa_3 = 8n$, $\kappa_4 = 48n$, etc., so that $c = \sqrt{\frac{2}{n}}$, $\gamma_1 = 2c$, $\gamma_2 = 6c^2$. Thus if we compare the χ^2 distribution with a cubed-normal distribution of the same coefficient of variation, we find that they are equally asymmetrical, but that the χ^2 distribution has a γ_2 which is $\frac{27}{25}$th that of the cubed normal. Hence the same transformation will nearly-abolish both κ_3 and κ_4 for both distributions.

The success of Wilson and Hilferty's transformation suggests strongly that we may use equations (1), (5) or (3) for the approximate normalization of moderately skew variates. This is an urgent problem in several applications of statistics to biology (Haldane, 1939). We evaluate a number whose mean and standard deviation in the case of random sampling are known. We desire to know whether it differs significantly from the mean. But the sampling distribution is found to be skew. If we can approximately normalize it, our tests of significance become far sharper. This problem is taken up in detail elsewhere. If $\gamma_1^2 > \frac{9}{16}\gamma_2$, and both are small, so that k^2 can be neglected, we can find m, n and k so that $(m+k^{\frac{1}{2}}x)^n$ has a given γ_1 and γ_2. By equations (3) and (4),

$$n = 1+\frac{4\gamma_1^2}{16\gamma_1^2-9\gamma_2}, \quad k = \frac{\gamma_1^2}{9(n-1)^2}, \quad m^{2n} = \frac{\kappa_2}{n^2k}.$$

If, however, γ_1 is not small, it is necessary to take several terms of equations (3).

If in an observed distribution of weights or volumes, the estimate of γ_1/c, or of $\kappa_1\kappa_3/\kappa_2^2$ is approximately 2, it will be reasonable to try whether γ_2/c^2 or $\kappa_1^2\kappa_4/\kappa_2^2$ approximates to $\frac{56}{9}$, and if so to try to fit a normal distribution to the cube roots of the variate. It will be seen that this does not imply that all the objects considered are of the same shape. On the contrary, such a distribution is to be expected whenever three mutually perpendicular measurements are normally distributed, provided that their coefficients of variation and correlation are not very different. And even the greatest differences in the latter, provided they are not negative, will have little effect.

Unfortunately, reliable estimates of γ_1 and γ_2 can only be obtained from samples of the order of 1000 or more. Rendel (in a paper to be published shortly) obtained the following estimates for the cumulants of the weight distribution of 1202 viable duck eggs, corrected for grouping. The unit is a gram:

$$k_1 = 73{\cdot}294, \quad k_2 = 42{\cdot}69, \quad k_3 = +118{\cdot}658, \quad k_4 = +937{\cdot}21.$$

Hence the estimates of c, γ_1 and γ_2 are

$$c = 0{\cdot}0887 \pm 0{\cdot}0018, \quad g_1 = +0{\cdot}432 \pm 0{\cdot}071, \quad g_2 = +0{\cdot}524 \pm 0{\cdot}142.$$

The standard errors are those appropriate to a normal distribution, but the true values cannot be very different. It is clear that g_1 and g_2 are significantly greater than the values of 0·27 and 0·13 which we should expect were the logarithms of weights normally distributed. They differ still more from what would be expected were the cube roots of weights normally distributed, or on any other hypothesis leading to a similar distribution.

Pearl (1905) lists the moments of the distributions of brain weights of eight European populations of 197 to 529 individuals. The various estimates of c are all close to 0·080, ranging from 0·074 to 0·083. Those of γ_1 range from $+0{\cdot}11$ to $+0{\cdot}40$, those of γ_2 from $-0{\cdot}30$ to $+1{\cdot}5$. The weighted means are

$$c = 0{\cdot}07965 \pm 0{\cdot}00106, \quad g_1 = +0{\cdot}2306 \pm 0{\cdot}0461, \quad g_2 = +0{\cdot}2661 \pm 0{\cdot}0922.$$

If the cube roots are normally distributed, we should expect $g_1 = +0{\cdot}16$, $g_2 = +0{\cdot}037$; if their logarithms are normally distributed, we should expect $g_1 = +0{\cdot}24$, $g_2 = +0{\cdot}10$. The latter distribution gives the better fit, but the first is not impossible.

Sinnott (1937) gives a graph of the distribution of the weights of squash fruits in an F_2, which is positively skew. He shows that a graph of the distribution of their logarithms, though negatively skew, is more symmetrical. There is a suggestion that a graph of the distribution of their cube roots would be even more so. Unfortunately the actual figures are not given, and since curve-fitting by eye is notoriously uncertain, no more can be said. It is much to be desired that, when the full data are not given, estimates of the first four moments or cumulants should be published.

11. SUMMARY

The first four moments of a number of powers and products of normal variates are calculated, with special reference to the probable distributions of weights, volumes, or areas of organs and organisms. In each case the first two measures (γ_1 and γ_2 or β_1 and β_2) of deviation from normality are obtained in terms of the coefficients of variation. The expressions obtained are almost independent of the correlation between the linear measurements, provided the coefficient of variation of the latter are approximately equal. The distribution found is perhaps applicable to data on brain weights, but not to data on ducks' eggs.

REFERENCES

CRAIG, C. C. (1936). On the frequency function of xy. *Ann. Math. Statist.* **7**, 1.

HALDANE, J. B. S. (1938). The approximate normalization of a class of frequency distributions. *Biometrika*, **29**, 392–404.

—— (1939). New data on partial sex-linkage in man. *Proc. Int. Genet. Congr.* p. 137.

MACALISTER, D. (1879). The law of the geometric mean. *Proc. Roy. Soc.* **24**, 367.

PEARL, R. (1905). Biometrical studies on man. I. Variation and correlation in brain-weight. *Biometrika*, **4**, 13–104.

PEARSON, K. (1905). Das Fehlergesetz und seine Verallgemeinerungen durch Fechner und Pearson. A rejoinder. *Biometrika*, **4**, 169–212.

SINNOTT, E. W. (1937). The relation of gene to character in quantitative inheritance. *Proc. Nat. Acad. Sci., Wash.*, **23**, 224–7.

WILSON, E. B. & HILFERTY, M. M. (1931). The distribution of Chi-square. *Proc. Nat. Acad. Sci., Wash.*, **17**, 684.

THE TRANSFORMATION OF DATA FROM ENTOMOLOGICAL FIELD EXPERIMENTS SO THAT THE ANALYSIS OF VARIANCE BECOMES APPLICABLE†

By GEOFFREY BEALL
Dominion Entomological Laboratory, Chatham, Ontario, Canada

1. Introductory

The present paper deals with experiments on the control of insects in the field. In such experimental work the problem to be investigated is whether more insects survive on plots which have been subjected to one treatment than on plots subjected to another. It will be shown in the present paper that the numbers of insects found per plot must vary in such a way that one cannot, strictly, subject the results to the analysis of variance, and it is proposed to find how the data may be transformed so that analysis of variance becomes applicable. Such transformation has been discussed by Bartlett (1936*a*, *b*) in connexion with entomological experiments, and by Tippett (1934) in connexion with industrial experiments.

2. Experimental results considered

The data used in the following work are results from seven insecticidal experiments arranged by the author at Chatham, Ontario. The work was carried out with replicated blocks containing plots subjected to treatments of which the assignment was random. This procedure, normal in agronomic work, was supplemented by one repetition of each treatment within a block. The assignment of the repetition of a treatment was independent of the first for that treatment, except that, of course, the same plot could not be chosen twice. This repetition was carried out to obtain estimates of variability within blocks. In these experiments complete counts were not made but random sampling was employed. Experiments on *Pyrausta nubilalis* Hubn., reported by Beall *et al.* (1939), for which results are shown in Tables 1 and 2, were made on one area at two different periods, whereas experiments on *Leptinotarsa decemlineata* Say, for which results are indicated in Tables 3 and 4, were carried out on contiguous areas at the same time. Three similar experiments were carried out in one place on the tobacco hornworm, *Phlegethontius quinquemaculata* Haw., for which the data are shown in Tables 5–7. Reference is also made to the data from a uniformity trial on insects of Beall (1939).

† Publication No. 2191, Division of Entomology, Science Service, Department of Agriculture, Ottawa, Canada.

Table 1. *Numbers of an insect,* Pyrausta nubilalis, *per plot. Experiment I*

Treatment	Block									
	1	2	3	4	5	6	7	8	9	10
1	15	23	21	31	22	14	18	11	21	34
1	27	20	23	33	34	27	17	13	20	26
2	19	12	34	16	20	10	24	23	14	13
2	11	28	37	16	26	18	19	13	10	9
3	16	15	22	25	13	21	18	38	27	10
3	19	16	18	21	19	24	21	18	12	18
4	14	23	10	19	17	18	7	18	8	17
4	34	21	9	34	19	9	15	16	12	12
5	16	16	19	26	15	11	18	23	27	6
5	23	12	12	10	12	17	13	21	12	9
6	12	14	17	10	14	24	24	17	7	3
6	15	16	15	28	13	22	11	7	5	4
7	43	28	35	36	50	69	62	63	42	40
7	47	81	30	69	35	29	71	47	50	43

Table 2. *Numbers of an insect,* Pyrausta nubilalis, *per plot. Experiment II*

Treatment	Block									
	1	2	3	4	5	6	7	8	9	10
1	32	38	27	7	13	14	26	25	22	30
1	18	40	39	12	19	26	30	19	18	28
2	6	23	8	4	3	18	26	27	17	19
2	9	14	20	13	15	14	15	19	19	10
3	10	21	25	10	13	20	33	48	28	27
3	4	21	26	4	9	14	30	18	27	18
4	2	17	11	3	10	10	26	13	22	17
4	24	13	13	10	6	14	28	11	34	7
5	13	2	5	0	18	10	33	23	20	34
5	17	22	23	8	14	16	26	22	15	34
6	13	10	21	4	10	8	17	15	13	16
6	17	9	29	5	18	15	19	16	27	23
7	37	58	28	11	24	44	30	44	56	45
7	44	71	55	20	26	27	43	52	39	58

Table 3. *Numbers of an insect,* Leptinotarsa decemlineata, *per plot. Experiment III*

Treatment	Block						
	1	2	3	4	5	6	7
1	305	391	420	355	287	175	454
1	207	364	639	527	293	248	397
2	97	49	21	12	3	10	10
2	93	51	25	37	4	12	1
3	270	105	341	469	82	57	221
3	153	190	348	212	100	285	309
4	7	42	34	8	1	10	4
4	12	2	22	4	1	3	3

Table 4. *Numbers of an insect,* Leptinotarsa decemlineata, *per plot. Experiment IV*

Treatment	Block					
	1	2	3	4	5	6
1	253	145	309	665	99	93
1	239	265	166	230	302	237
2	16	13	74	110	14	5
2	95	54	159	108	14	13
3	18	130	165	137	153	78
3	40	137	118	142	239	63
4	2	0	22	6	129	3
4	2	1	31	8	9	8

Table 5. *Numbers of an insect,* Phlegethontius quinquemaculata, *per plot. Experiment V*

Treatment	Block					
	1	2	3	4	5	6
1	6	5	6	13	6	11
1	4	15	13	6	10	15
2	0	1	1	0	1	1
2	2	2	1	4	1	1
3	15	17	22	28	8	16
3	12	22	16	11	13	25

Table 6. *Numbers of an insect,* Phlegethontius quinquemaculata, *per plot. Experiment VI*

Treatment	Block					
	1	2	3	4	5	6
1	12	13	9	4	11	4
1	13	9	5	7	5	10
2	13	6	8	1	5	7
2	20	9	9	4	12	7
3	7	9	5	4	8	9
3	7	9	4	7	3	2
4	1	1	0	1	4	3
4	2	2	2	1	4	5
5	13	7	12	3	6	11
5	11	5	4	1	9	8
6	7	6	8	9	6	5
6	8	10	2	4	4	12

Table 7. *Numbers of an insect,* Phlegethontius quinquemaculata, *per plot. Experiment VII*

Treatment	Block					
	1	2	3	4	5	6
1	10	20	14	10	17	14
1	7	14	12	23	20	13
2	11	21	16	17	19	7
2	17	11	14	17	21	13
3	0	7	3	2	3	1
3	1	2	1	1	0	4
4	3	12	4	5	5	2
4	5	6	3	5	5	4
5	3	3	3	1	3	6
5	5	5	6	1	2	4
6	11	15	15	13	26	24
6	9	22	16	10	26	13

3. The relationship between the standard deviation and the mean in the experimental data

If x is the number of insects on one of a group of small contiguous areas, say plots, within a larger area, say a block, let the expectation of x over all these plots be M and the standard deviation be σ; then over a number of the larger

areas, when the insects are distributed in a completely random fashion, from the Poisson distribution,

$$\sigma^2 = M. \tag{1}$$

As is discussed by 'Student' (1919) one cannot, however, anticipate that (1) will be satisfied when organisms occur in groups, as, say, when insects come from masses of eggs, or when there is a change in expectation from plot to plot within a block. Generally, σ^2 will tend to be greater than M and we can only say

$$\sigma^2 = f(M). \tag{2}$$

The form of $f(M)$, in (2), must be considered carefully, since it bears on the form of the transformation which may be developed to make the standard deviation independent of the mean.

In dealing with (2), Bartlett (1936*a*) started by supposing that, approximately,

$$\sigma^2 = KM, \tag{3}$$

where K is a constant. Generally, in field data, however, the relationship between σ^2 and M, or of their respective estimates, s^2 and $\bar{x}$, does not, as in Fig. 1, appear to be linear; rather, the departure of s^2 from $\bar{x}$ becomes disproportionately great as $\bar{x}$ increases. This relationship between departures and the magnitude of the mean has been discussed by Clapham (1936) in connexion with data on the distribution of organisms differing from insects as much as flowering plants, and he showed that only those distributions with very low mean have the squared standard deviation close to the mean.

Our discussion above on the shortcomings of (3) suggests the conclusion that

$$\sigma^2 - M \propto M \tag{4}$$

is generally untrue. We propose to consider the possibility that the curvilinearity of (2) might be better met by supposing that

$$\sigma^2 - M \propto M^2. \tag{5}$$

Equation (5) leads to

$$\sigma^2 = M + kM^2, \tag{6}$$

where k is a constant. It will be noticed that

$$k = (\sigma^2 - M)\, M^{-2} \tag{7}$$

is the Charlier coefficient of disturbance from a Poisson distribution. This coefficient was employed by Beall (1935).

It is possible to consider the suitability of (3), as compared with (6), by finding how, respectively, they fit observations on s^2 and $\bar{x}$. To fit exactly is difficult, and it was found necessary to fall back on an empirical determination of K and of k; thus, if there are a number of pairs of estimates, $\bar{x}$ and s^2, from (3) and (6) we estimate

$$K = \Sigma s^2 / \Sigma \bar{x}, \tag{8}$$

$$k = (\Sigma s^2 - \Sigma \bar{x}) / \Sigma \bar{x}^2, \tag{9}$$

where Σ represents the summation over all pairs.

Since in the work presented in § 2, $\bar{x}$ and s^2, being based on only two observations, are highly variable, these experiments do not show clearly the suitability of (3) and (6). Accordingly, reference is made instead to the data from the uniformity trial on *Leptinotarsa decemlineata* Say of Beall (1939). When the mean and standard deviation of 144 sampling units within each of 16 areas were considered, the estimates from (8) and (9) were $K = 2{\cdot}405$ and $k = 0{\cdot}2548$. For these values from (1), (3) and (6), curves, described as lines 1, 2 and 3 respectively, are plotted in Fig. 1, the observed values of mean and squared standard deviation

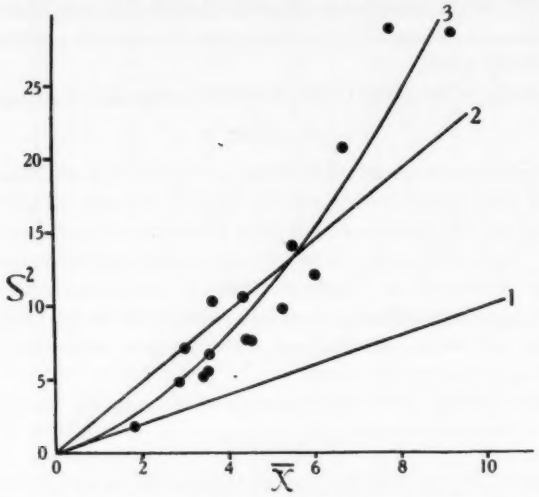


Fig. 1. The squared standard deviation plotted against the mean for 144 small areas within each of 16 large areas; line 1 is from equation (1), line 2 from (3) and line 3 from (6). The counts had been made on *Leptinotarsa decemlineata* Say.

are also shown. In the cases where the mean is near unity the departure of the squared standard deviation from the mean, i.e. from line 1, appears to be trivial, but as the mean increases the departure becomes more marked. It can be seen that the observations lie more snugly about line 3 from (6) than about line 2 from (3). Generally, for the data from field studies the same effect has been observed. Such results suggest that (6) may be generally a better approximation to the form of $f(M)$ than (3) and make it preferable to proceed with the analysis of data from the assumption (6).

4. The transformations of field data

Fig. 1 shows clearly how, within an area, the variability of the numbers of insects on sub-areas is related to the mean number of insects per sub-area. This relationship will make invalid the use of the analysis of variance on experimental

results involving counts on insects, since the expectation of the variance should be the same for all plots. To overcome this invalidity, Bartlett (1936*a*) suggested transforming the observations, x, from the basis of (3). The transformation found was $x^{\frac{1}{2}}$, which Bartlett modified to $(x+\frac{1}{2})^{\frac{1}{2}}$. From § 3 it was seen, however, that for field data the relationship between standard deviation and mean may be represented better by equation (6) than by (3), and, since the form of the transformation depends on the form of $f(M)$, a fresh transformation must be sought. A transformation, as is developed in the Appendix to the present paper, is suggested by the method of Tippett (1934), i.e.

$$x' = k^{-\frac{1}{2}} \sinh^{-1}(kx)^{\frac{1}{2}}. \tag{10}$$

An advance note of this transformation was published by Beall (1940). The adequacy of this transformation must be judged from the extent to which it stabilizes variability. In (10), if we express $\sinh^{-1}(kx)^{\frac{1}{2}}$, when $kx < 1$, as a well-known series, we have

$$x' = x^{\frac{1}{2}} - \tfrac{1}{6} kx^{\frac{1}{2}} + \tfrac{3}{40} k^2 x^{\frac{1}{2}} - \tfrac{5}{112} k^3 x^{\frac{1}{2}} + \ldots, \tag{11}$$

where it is obvious that for $k = 0$, $x' = x^{\frac{1}{2}}$. Of course, for large values of kx, x' varies almost as $\log x$, or as the $\log(x+1)$ used by Williams (1937), and so our proposed expansion may be regarded, for practical purposes, as embracing the root and logarithmic transformations.

Table 8 gives the transformation, (10), for a probable range of observations, x, and for k at intervals which will probably be close enough for practical purposes. This table was computed in part by inverse interpolation from the table of hyperbolic functions of the *Smithsonian Mathematical Tables* (Becker & Van Orstrand, 1931), and in part from (12). Should values of x' be required outside those of Table 8, these can conveniently be calculated from

$$x' = k^{-\frac{1}{2}} \log_e \{(kx)^{\frac{1}{2}} + (1+kx)^{\frac{1}{2}}\}. \tag{12}$$

In preparing Table 8 the question arose of whether, instead of dealing with $k^{-\frac{1}{2}} \sinh^{-1}(kx)^{\frac{1}{2}}$, one should not use $k^{-\frac{1}{2}} \sinh^{-1}\{k^{\frac{1}{2}}(x+\frac{1}{2})^{\frac{1}{2}}\}$ in the same way as Bartlett (1936*a*) dealt with the transformation, $(x+\frac{1}{2})^{\frac{1}{2}}$, instead of $x^{\frac{1}{2}}$. This modification was rejected on the basis of results of the transformation, as discussed in § 5, since it was found that the addition of $\frac{1}{2}$ made little difference and did not give, consistently, an improvement.

For field data, in making the transformation (10), it is necessary to estimate the value of k empirically by (9) for which estimates $\bar{x}$, of the mean and s, of the standard deviation, must be found. The most obvious method in practice of making these estimates seems to be to put more than one plot subjected to a given treatment in a block and so to estimate the chance variation of results for a plot within a block. In the present work, as is discussed in § 2, two plots were subjected to a given treatment in each block and this is probably good practice.

Table 8. *The transformation* $x' = k^{-\frac{1}{2}} \sinh^{-1} (kx)^{\frac{1}{2}}$

x \ k	0·00	0·02	0·04	0·06	0·08	0·10	0·15	0·20	0·25	0·30	0·40	0·50	0·60	0·80	1·00
0	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00
1	1·00	1·00	0·99	0·99	0·99	0·98	0·98	0·97	0·96	0·96	0·94	0·93	0·92	0·90	0·88
2	1·41	1·40	1·40	1·39	1·38	1·37	1·35	1·33	1·32	1·30	1·27	1·25	1·22	1·18	1·15
3	1·73	1·72	1·70	1·68	1·67	1·66	1·62	1·59	1·57	1·54	1·50	1·46	1·42	1·37	1·32
4	2·00	1·97	1·95	1·93	1·91	1·89	1·84	1·80	1·76	1·73	1·67	1·62	1·58	1·50	1·44
5	2·24	2·20	2·17	2·14	2·11	2·08	2·02	1·97	1·92	1·88	1·81	1·75	1·70	1·61	1·54
6	2·45	2·40	2·36	2·32	2·29	2·25	2·18	2·12	2·06	2·01	1·93	1·86	1·80	1·71	1·63
7	2·65	2·59	2·54	2·49	2·45	2·41	2·32	2·25	2·18	2·13	2·04	1·96	1·89	1·78	1·70
8	2·83	2·76	2·70	2·64	2·59	2·54	2·45	2·36	2·29	2·23	2·13	2·04	1·97	1·85	1·76
9	3·00	2·92	2·84	2·78	2·72	2·67	2·56	2·47	2·39	2·32	2·21	2·12	2·04	1·92	1·82
10	3·16	3·07	2·98	2·91	2·85	2·79	2·66	2·56	2·48	2·40	2·28	2·18	2·10	1·97	1·87
11	3·32	3·21	3·11	3·03	2·96	2·89	2·76	2·65	2·56	2·48	2·35	2·25	2·16	2·02	1·91
12	3·46	3·34	3·23	3·14	3·07	3·00	2·85	2·73	2·63	2·55	2·41	2·30	2·21	2·07	1·96
13	3·61	3·47	3·35	3·25	3·17	3·09	2·93	2·81	2·70	2·62	2·47	2·36	2·26	2·11	1·99
14	3·74	3·59	3·46	3·35	3·26	3·18	3·01	2·88	2·77	2·68	2·52	2·40	2·31	2·15	2·03
15	3·87	3·70	3·56	3·45	3·35	3·26	3·08	2·94	2·83	2·73	2·57	2·45	2·35	2·19	2·06
16	4·00	3·81	3·66	3·54	3·43	3·34	3·15	3·01	2·89	2·79	2·62	2·49	2·39	2·22	2·09
17	4·12	3·92	3·76	3·63	3·51	3·42	3·22	3·07	2·94	2·84	2·67	2·53	2·42	2·25	2·12
18	4·24	4·02	3·85	3·71	3·59	3·49	3·28	3·12	2·99	2·88	2·71	2·57	2·46	2·28	2·15
19	4·36	4·12	3·94	3·79	3·67	3·56	3·34	3·18	3·04	2·93	2·75	2·61	2·49	2·31	2·18
20	4·47	4·22	4·02	3·87	3·74	3·62	3·40	3·23	3·09	2·97	2·79	2·64	2·52	2·34	2·20
21	4·58	4·31	4·11	3·94	3·80	3·69	3·46	3·28	3·13	3·01	2·82	2·68	2·56	2·37	2·23
22	4·69	4·40	4·18	4·01	3·87	3·75	3·51	3·32	3·18	3·05	2·86	2·71	2·58	2·39	2·25
23	4·80	4·49	4·26	4·08	3·93	3·81	3·56	3·37	3·22	3·09	2·89	2·74	2·61	2·42	2·27
24	4·90	4·57	4·34	4·15	3·99	3·86	3·61	3·41	3·26	3·13	2·92	2·77	2·64	2·44	2·29
25	5·00	4·66	4·41	4·21	4·05	3·92	3·65	3·45	3·29	3·16	2·95	2·79	2·66	2·46	2·31
26	5·10	4·74	4·48	4·27	4·11	3·97	3·70	3·49	3·33	3·20	2·98	2·82	2·69	2·48	2·33
27	5·20	4·82	4·54	4·33	4·16	4·02	3·74	3·53	3·37	3·23	3·01	2·85	2·71	2·51	2·35
28	5·29	4·89	4·61	4·39	4·22	4·07	3·78	3·57	3·40	3·26	3·04	2·87	2·73	2·53	2·37
29	5·39	4·97	4·67	4·45	4·27	4·12	3·82	3·61	3·43	3·29	3·07	2·89	2·76	2·54	2·39
30	5·48	5·04	4·74	4·51	4·32	4·16	3·86	3·64	3·46	3·32	3·09	2·92	2·78	2·56	2·40
31	5·57	5·11	4·80	4·56	4·37	4·21	3·90	3·67	3·50	3·35	3·12	2·94	2·80	2·58	2·42
32	5·66	5·18	4·86	4·61	4·42	4·25	3·94	3·71	3·53	3·38	3·14	2·96	2·82	2·60	2·43
33	5·74	5·25	4·91	4·66	4·46	4·30	3·98	3·74	3·55	3·40	3·16	2·98	2·84	2·62	2·45
34	5·83	5·32	4·97	4·71	4·51	4·34	4·01	3·77	3·58	3·43	3·19	3·00	2·86	2·63	2·46
35	5·92	5·38	5·03	4·76	4·55	4·38	4·05	3·80	3·61	3·45	3·21	3·02	2·88	2·65	2·48
36	6·00	5·45	5·08	4·81	4·59	4·42	4·08	3·83	3·64	3·48	3·23	3·04	2·89	2·66	2·49
37	6·08	5·51	5·13	4·85	4·64	4·46	4·11	3·86	3·66	3·50	3·25	3·06	2·91	2·68	2·51
38	6·16	5·57	5·18	4·90	4·68	4·49	4·14	3·89	3·69	3·53	3·27	3·08	2·93	2·69	2·52
39	6·24	5·63	5·23	4·94	4·72	4·53	4·17	3·91	3·71	3·55	3·29	3·10	2·94	2·71	2·53

38	6·16	5·57	5·18	4·90	4·68	4·49	4·14	3·89	3·69	3·53	3·27	3·08	2·93	2·69	2·52
39	6·24	5·63	5·23	4·94	4·72	4·53	4·17	3·91	3·71	3·55	3·29	3·10	2·94	2·71	2·53
40	6·32	5·69	5·28	4·99	4·76	4·57	4·20	3·94	3·74	3·57	3·31	3·12	2·96	2·72	2·54
41	6·40	5·75	5·33	5·03	4·79	4·60	4·23	3·97	3·76	3·59	3·33	3·13	2·98	2·73	2·56
42	6·48	5·81	5·38	5·07	4·83	4·63	4·26	3·99	3·78	3·61	3·35	3·15	2·99	2·75	2·57
43	6·56	5·86	5·43	5·11	4·87	4·67	4·29	4·02	3·81	3·63	3·37	3·17	3·01	2·76	2·58
44	6·63	5·92	5·47	5·15	4·90	4·70	4·32	4·04	3·83	3·65	3·39	3·18	3·02	2·77	2·59
45	6·71	5·97	5·52	5·19	4·94	4·73	4·35	4·07	3·85	3·67	3·40	3·20	3·03	2·79	2·60
46	6·78	6·03	5·56	5·23	4·97	4·76	4·37	4·09	3·87	3·69	3·42	3·21	3·05	2·80	2·61
47	6·86	6·08	5·61	5·27	5·01	4·80	4·40	4·11	3·89	3·71	3·44	3·23	3·06	2·81	2·62
48	6·93	6·13	5·65	5·30	5·04	4·83	4·42	4·13	3·91	3·73	3·45	3·24	3·08	2·82	2·63
49	7·00	6·18	5·69	5·34	5·07	4·85	4·45	4·16	3·93	3·75	3·47	3·26	3·09	2·83	2·64
50	7·07	6·23	5·73	5·38	5·10	4·88	4·47	4·18	3·95	3·77	3·48	3·27	3·10	2·84	2·65
55	7·42	6·47	5·93	5·55	5·26	5·02	4·59	4·28	4·04	3·85	3·56	3·34	3·16	2·90	2·70
60	7·75	6·70	6·11	5·70	5·39	5·15	4·70	4·37	4·13	3·93	3·62	3·40	3·22	2·94	2·74
65	8·06	6·91	6·28	5·85	5·52	5·27	4·79	4·46	4·20	4·00	3·69	3·45	3·27	2·99	2·78
70	8·37	7·11	6·44	5·98	5·64	5·38	4·88	4·54	4·28	4·07	3·74	3·50	3·32	3·03	2·82
75	8·66	7·30	6·58	6·11	5·76	5·48	4·97	4·61	4·34	4·13	3·80	3·55	3·36	3·07	2·86
80	8·94	7·47	6·73	6·23	5·86	5·57	5·05	4·68	4·41	4·19	3·85	3·60	3·40	3·10	2·89
85	9·22	7·64	6·86	6·34	5·96	5·66	5·13	4·75	4·47	4·24	3·90	3·64	3·44	3·14	2·92
90	9·49	7·80	6·98	6·45	6·06	5·75	5·20	4·81	4·52	4·29	3·94	3·68	3·48	3·17	2·95
95	9·75	7·96	7·10	6·55	6·15	5·83	5·26	4·87	4·57	4·34	3·98	3·72	3·51	3·20	2·97
100	10·00	8·10	7·22	6·65	6·23	5·91	5·33	4·93	4·62	4·39	4·02	3·75	3·54	3·23	3·00
110	10·49	8·38	7·43	6·83	6·39	6·05	5·45	5·03	4·72	4·47	4·10	3·82	3·60	3·28	3·05
120	10·95	8·64	7·63	6·99	6·54	6·18	5·56	5·13	4·80	4·55	4·16	3·88	3·66	3·33	3·09
130	11·40	8·88	7·81	7·15	6·67	6·31	5·66	5·21	4·88	4·62	4·23	3·94	3·71	3·37	3·13
140	11·83	9·10	7·98	7·29	6·80	6·42	5·75	5·30	4·96	4·69	4·29	3·99	3·76	3·42	3·17
150	12·25	9·31	8·14	7·42	6·91	6·53	5·84	5·37	5·02	4·75	4·34	4·04	3·80	3·45	3·20
160	12·65	9·51	8·29	7·55	7·02	6·62	5·92	5·44	5·09	4·81	4·39	4·08	3·84	3·49	3·23
170	13·04	9·70	8·43	7·67	7·13	6·72	6·00	5·51	5·15	4·86	4·44	4·13	3·88	3·52	3·26
180	13·42	9·88	8·57	7·78	7·23	6·81	6·07	5·57	5·20	4·92	4·48	4·17	3·92	3·56	3·29
190	13·78	10·05	8·69	7·88	7·32	6·89	6·14	5·63	5·26	4·96	4·52	4·20	3·95	3·59	3·32
200	14·14	10·21	8·81	7·98	7·41	6·97	6·20	5·69	5·31	5·01	4·57	4·24	3·99	3·61	3·34
210	14·49	10·36	8·93	8·08	7·49	7·04	6·26	5·74	5·36	5·05	4·60	4·27	4·02	3·64	3·37
220	14·83	10·51	9·04	8·17	7·57	7·11	6·32	5·79	5·40	5·10	4·64	4·31	4·05	3·67	3·39
230	15·17	10·65	9·14	8·26	7·65	7·18	6·38	5·84	5·45	5·14	4·68	4·34	4·08	3·69	3·41
240	15·49	10·79	9·25	8·34	7·72	7·25	6·43	5·89	5·49	5·18	4·71	4·37	4·11	3·72	3·43
250	15·81	10·92	9·34	8·42	7·79	7·31	6·49	5·93	5·53	5·21	4·74	4·40	4·13	3·74	3·45
260	16·12	11·05	9·44	8·50	7·86	7·37	6·54	5·98	5·57	5·25	4·77	4·42	4·16	3·76	3·47
270	16·43	11·17	9·53	8·58	7·92	7·43	6·58	6·02	5·61	5·28	4·80	4·45	4·18	3·78	3·49
280	16·73	11·29	9·61	8·65	7·99	7·49	6·63	6·06	5·64	5·32	4·83	4·48	4·20	3·80	3·51
290	17·03	11·40	9·70	8·72	8·05	7·54	6·68	6·10	5·68	5·35	4·86	4·50	4·23	3·82	3·53
300	17·32	11·51	9·78	8·79	8·10	7·60	6·72	6·14	5·71	5·38	4·88	4·53	4·25	3·84	3·55

In the special case where there are two plots for the ith treatment ($i = 1, \ldots, n$) in the jth block ($j = 1, \ldots, N$), and so two observations, x_{ij1} and x_{ij2}, the estimate of the mean will be written $x_{ij.}$ and of the squared standard deviation

$$s^2_{ij} = \tfrac{1}{2}(x_{ij1} - x_{ij2})^2. \tag{13}$$

Then from (9), we estimate k from

$$k = 2\left\{\sum_{i=1}^{n}\sum_{j=1}^{N}(x_{ij1} - x_{ij2})^2 - \sum_{i=1}^{n}\sum_{j=1}^{N}(x_{ij1} + x_{ij2})\right\}\left\{\sum_{i=1}^{n}\sum_{j=1}^{N}(x_{ij1} + x_{ij2})^2\right\}^{-1} \tag{14}$$

and the calculation is very light.

5. Results showing the effect of the transformation on the variability of data

The adequacy of our proposed transformation may be judged in two ways: first, with respect to its effect, which we shall consider in the present section, on the differences between repetitions of a treatment within a block, and secondly, with respect to its effect, which we shall consider in § 6, on the behaviour of the quantities submitted to the analysis of variance.

It is a fundamental assumption in the analysis of variance that the chance variability for each plot shall be, when the effect of block and of treatment are removed, normally distributed with a standard deviation common to all plots, in which situation of course the standard deviation of the chance variability for a given plot is independent of the expectation for that plot. In the data of the present work, where each treatment is repeated in each block, it is possible to examine the estimates of this standard deviation, s_{ij}, and of the expectation, $x_{ij.}$. For a clear graphical illustration of the situation consider Fig. 2, as obtained from the original data of Experiment III on *Leptinotarsa decemlineata*, where s_{ij} is plotted against $x_{ij.}$ and contrast this situation with that obtaining for the corresponding quantities s'_{ij} and $x'_{ij.}$ obtained after transformation ($k = 0{\cdot}08$) in Fig. 3.

In Fig. 2 the points are widely scattered as is natural from a sample of two; nevertheless, it is apparent that for the smallest values of $x_{ij.}$ the values of s_{ij} are correspondingly small and fall in a close group. In Fig. 3 the cluster of observations in the lower left-hand corner of the previous diagram has disappeared, and generally the scatter appears to be independent of $x'_{ij.}$, so that apparently the transformation gave satisfactory results. The nature of the material involved is such that it does not seem possible to examine the relationship under consideration more exactly, nor to summarize exactly the corresponding results for the other treatments; it can only be said that the same type of result appeared although the magnitude of the relationship before transformation depended on the magnitude of the differences between the effects of treatments.

The results shown in Figs. 2 and 3 suggest that the proposed transformation has tended to make the standard deviation independent of the mean, in accordance

with the assumptions underlying the analysis of variance. In using this procedure one actually assumes, more broadly, that a common standard deviation exists, so that the homoscedasticity of observations before and after transformation should be tested. Thus it is assumed that x_{ij1} and x_{ij2} are observations from a normal

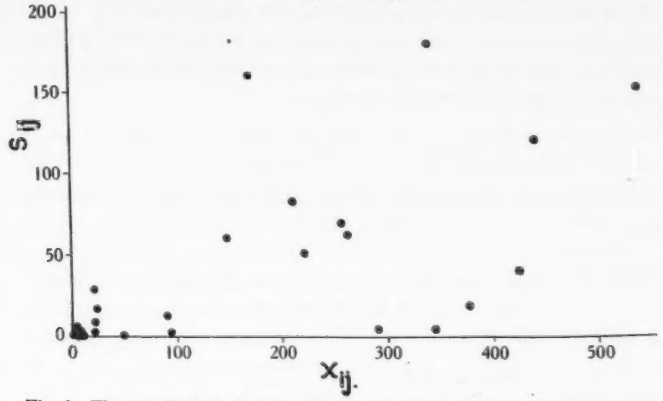


Fig. 2. The standard deviation and mean as estimated from plots by pairs, with untransformed data on *Leptinotarsa decemlineata*.

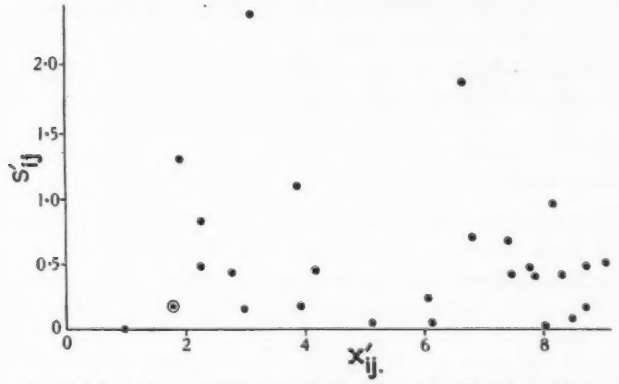


Fig. 3. The standard deviation and mean as estimated from plots by pairs, with the transformed data on *Leptinotarsa decemlineata* Say, i.e. using $x' = k^{-1} \sinh^{-1} (kx)^{\frac{1}{2}}$, $(k = 0{\cdot}08)$.

population with a standard deviation, σ, which is independent of i and j. Then $\left\{\sum_{k=1}^{2} (x_{ijk} - x_{ij.})^2\right\} \frac{1}{\sigma^2}$ is distributed as χ^2 with one degree of freedom.† Accordingly,

† In the L_1 test, discussed by Nayer (1936), this case of estimates of standard deviation with one degree of freedom is troublesome since zero values tend to arise when dealing with grouped or integral observations. When this is the case L_1, which is the ratio of an arithmetic to a geometric mean of sums of squares, cannot be calculated. The present treatment may therefore have a wider application.

$y_{ij} = (x_{ij1} - x_{ij2})/\sqrt{2}\sigma$ should be distributed normally with unit standard deviation for all i and j. In order to test the hypothesis of normality with unit standard deviation it is only necessary to test for leptokurtosis; for the distribution must be symmetrical since the sign of differences, and therefore of y_{ij}, is a matter of chance. Since the number of items involved will almost certainly be < 100, and since the population mean is zero, the w_n criterion of Geary (1935) will provide an appropriate test. In using this criterion we must find the ratio of the mean deviation to the standard deviation, i.e.

$$w_n = \left\{ \sum_{i=1}^{n} \sum_{j=1}^{N} | x_{ij1} - x_{ij2} | \right\} \left\{ nN \sum_{i=1}^{n} \sum_{j=1}^{N} (x_{ij1} - x_{ij2})^2 \right\}^{-\frac{1}{2}}. \qquad (15)$$

Of course, values of w_n may be calculated for transformed data by substitution of x'_{ijk} for x_{ijk}.

Table 9. *The w_n test on the homoscedasticity of counts by plots within a block for six field experiments*

Experiment	nN	Upper 5% limit	Lower 5% limit	Untransformed		Transformed (Bartlett)		Value of k		Transformed (Beall)	
				w_n	Departure by S.D.	w_n	Departure by S.D.	Estimated	Employed	w_n	Departure by S.D.
I	70	0·841	0·757	0·6659	−5·33	0·7525	−1·91	0·078	0·08	0·7846	−0·64
II	70	0·841	0·757	0·7885	−0·49	0·7807	−0·79	0·046	0·04	0·7499	−2·01
III	28	0·866	0·737	0·5973	−5·33	0·6431	−4·15	0·084	0·08	0·6838	−3·11
IV	24	0·872	0·732	0·5692	−5·67	0·6948	−2·67	0·285	0·30	0·7370	−1·66
V	18	0·881	0·728	0·7554	−1·16	0·8155	+0·13	0·082	0·08	0·7823	−0·58
VI	36	0·857	0·745	0·7959	−0·22	0·8166	+0·38	0·019	0·02	0·8130	+0·28

Values of w_n from the untransformed observations and from the transformed observations, both following Bartlett (i.e. the transformation $(x+\frac{1}{2})^{\frac{1}{2}}$) and following the line suggested in the present paper, are shown in Table 9 for the field data of Tables 1–6. For the second transformation the values of k as calculated from (14) are shown as well as the nearest value of k entered in Table 8. For each experiment the value of nN and also the 0·05 limits of probability, from Geary (1935), are shown. There are also shown the departures of observed w_n from the expected value in terms of the standard deviation, a useful criterion since the distribution of w_n is almost normal. From Table 9 it can be seen that out of the three experiments in which w_n fell beyond the lower 5% limit of probability for the untransformed data and the data transformed as $(x+\frac{1}{2})^{\frac{1}{2}}$, in only one experiment did w_n fall so with the final transformation. The results for Experiment II, in which w_n is decreased by the transformation, are peculiar. Consideration of the departures from the mean in terms of the standard deviation indicates more clearly the improvement effected by each transformation

and how the transformation suggested in the present work secures an improvement of the same, but more marked, character than that secured from the transformation of Bartlett. The results suggest that while homoscedasticity may not be attained always, it will be approached by means of the proposed transformation.

6. THE EFFECT OF THE TRANSFORMATION ON THE ANALYSIS OF VARIANCE

As was indicated at the beginning of § 5, our proposed transformation besides making the variability within a block for a repeated treatment the same for all treatments and blocks, should also provide quantities satisfying the assumptions underlying the analysis of variance. Since it is not quite clear how, in so far as the transformation is satisfactory in the first way, it will necessarily be satisfactory in the second, it will be well to consider directly the suitability of our transformed values for the analysis of variance.

In the application of the analysis of variance one would deal with $x_{ij.}$ rather than with x_{ijk} and suppose that

$$x_{ij.} = A + B_i + C_j + D_{ij}, \tag{16}$$

where A is a contribution from the general level of population on the experimental area, B_i the contribution of the ith treatment and C_j the contribution of the jth block. The remainder term, D_{ij}, is called the interaction of treatments and blocks. Of course, the present discussion on the untransformed values, $x_{ij.}$, holds for the transformed values, $x'_{ij.} = \frac{1}{2}(x'_{ij1} + x'_{ij2})$ when the appropriate symbols, A', B'_i, C'_j and D'_{ij} are used.

In material satisfying the conditions underlying the analysis of variance, for the observations under each treatment, the calculated squared standard deviation is

$$s_i^2 = \frac{1}{N-1} \sum_{j=1}^{N} (x_{ij.} - x_{i..})^2. \tag{17}$$

Following the argument of the analysis of variance, $x_{ij.} - x_{i..}$, of which the mean is 0, is an estimate of $C_j + D_{ij}$, in which the two terms are independent; hence the expectation of s_i^2 is

$$\sigma_i^2 = \sigma_{C_j}^2 + \sigma_{D_{ij}}^2, \tag{18}$$

where σ_{Cj} and σ_{Dij} are the standard deviations of the parameters, C_j and D_{ij}, respectively, and are independent of treatment. Accordingly, s_i should be independent of treatment and distributed as an estimate of σ_i, having $N-1$ degrees of freedom. Conversely, if $x_{ij.}$ cannot be built of the independent terms of (16), then the various values of s_i will not be distributed as estimates of a single standard deviation. The hypothesis that the values of s_i in any one experiment are estimates of one quantity may be tested.†

† From correspondence with Dr R. W. B. Jackson, the writer has learned that he had arrived independently at the same test.

The results of the tests on the homogeneity of the values, s_i, within the six experiments treated in the present paper are presented in Table 10, where the value of the L_1 criterion is shown for the original data and for the transformed values together with the appropriate 0·05 and 0·01 levels of probability (Nayer, 1936). From Table 10 it can be seen that of the values of L_1 obtained from the original data, all but one are near or beyond the 0·05 level of significance, but that after transformation all are moved in to less significant values. Accordingly, the values of s_i, when calculated from the original data, appear heterogeneous but the corresponding values obtained after the transformation appear homogeneous. Thus it is more probable that the analysis of variance is applicable to the transformed data than to the untransformed.

Table 10. *The homogeneity, as measured by the criterion L_1, of the estimates s_i for various values of i before and after transformation*

	Experiment					
	1	2	3	4	5	6
L_1 before transformation	0·867	0·833	0·325	0·657	0·344	0·680
L_1 after transformation	0·864	0·941	0·766	0·688	0·813	0·730
1% limit	0·757	0·757	0·604	0·542	0·514	0·583
5% limit	0·812	0·812	0·707	0·656	0·648	0·673

In Table 10 we have tested the homogeneity of the estimates, s_i, as in § 5 we tested the homogeneity of s_{ij}, that is without reference to the values of the associated means. In view of our original assumptions we are, however, interested in the possibility that the standard deviations, as calculated, might show every sign of being estimates of a common standard deviation and yet be dependent on the associated means. Accordingly, we have investigated such dependence roughly by fitting by least squares a first order regression of s_i on $x_{i..}$. From this fitting we record the sign of the regression as follows:

	Experiment					
	1	2	3	4	5	6
Before transformation	+*	+**	+*	+**	+	+
After transformation	−	+	−	−*	+	+

By a single asterisk we have indicated cases where the reduction in variability effected by the regression passed the 5 % probability limit and by a double asterisk where it passed the 1 % limit. Several points may be noted. (1) In two cases (Experiments 2 and 6) after transformation the residual sum of squares about the regression was greater than the reduction in squares due to the regression, whereas it was consistently less before transformation. (2) As can be seen above, the regression generally did not effect a significant reduction in variability after transformation but did before (the small number of degrees of freedom made high significance difficult of attainment). (3) After transformation the sign of the regression seemed to be a chance matter, whereas before transformation it was consistently positive. These results suggest that the transformation proposed did tend to make the variability within a given treatment independent of the mean for that treatment.

7. THE EFFECT OF TRANSFORMATION UPON THE CONCLUSIONS FROM THE ANALYSIS OF VARIANCE

It has been shown in §§ 5 and 6 that the analysis of variance can be made on entomological data when a suitable transformation has been effected. It is of practical interest to see what numerical effect such transformation will have upon tests on the significance of, say, the effect of treatment and the significance of differences for treatments.

First, consider the numerical results to be obtained from the analysis of variance (1) without and (2) with transformation. Thus the mean square ascribable to blocks, treatments and their interaction is shown in Table 11, for six experiments of which the data are given in § 2; parallel results are presented for untransformed observations and for observations transformed by (10) with the values of k from Table 9. To facilitate the comparison of the results, the mean square for blocks and for treatments is expressed in terms of the estimate for interaction, as the F of Snedecor (1934), and presented in each case. The transformation of the data has modified the conclusions to be drawn from the analysis of variance in Table 11, in that there are considerable changes in the criterion, F, for treatments or for blocks. In the examples shown the effect of treatments was highly significant in all cases and so the changes introduced by transformation did not alter the conclusions, as would have been the case for less definite effects.

Consider next the effect of transformation on the significance of differences between the means for treatments as tested by the criterion, t, calculated with such estimates of mean square as the interaction of Table 11. For illustration, values of t, from the data on *Leptinotarsa decemlineata* (Experiment III), are shown in Table 12 for each possible comparison of treaments when untransformed data are used, when the transformation, $(x+\frac{1}{2})^{\frac{1}{2}}$, as suggested by Bartlett (1936*a*) is used, and when the transformation, $k^{-\frac{1}{2}}\sinh^{-1}(kx)^{\frac{1}{2}}$, as suggested in the present paper is used. In order that the influence of the level of population under each

Table 11. *The analysis of variance of untransformed and transformed data in six experiments*

Variation	Degrees of freedom	Untransformed data		Transformed data	
		Mean square	*F*	Mean square	*F*
Experiment I. ***P. nubilalis***					
Between blocks	9	92·8	1·21	0·565	1·66
Between treatments	6	2,839·0	36·95	7·51	22·03
Interaction	54	76·8	—	0·341	—
Experiment II. ***P. nubilalis***					
Between blocks	9	577·0	6·72	5·37	10·55
Between treatments	6	1,721·0	20·04	8·69	17·07
Interaction	54	85·9	—	0·509	—
Experiment III. ***L. decemlineata***					
Between blocks	6	20,172·0	2·26	4·67	3·03
Between treatments	3	390,932·0	43·77	111·5	72·16
Interaction	18	8,931·0	—	1·54	—
Experiment IV. ***L. decemlineata***					
Between blocks	5	12,960·0	2·06	2·06	1·96
Between treatments	3	124,054·0	20·40	20·40	19·41
Interaction	15	6,727·0	—	1·05	—
Experiment V. ***P. quinquemaculata***					
Between blocks	5	27·4	2·74	0·349	4·19
Between treatments	2	752·0	75·33	19·5	233·77
Interaction	10	9·98	—	0·083	—
Experiment VI. ***P. quinquemaculata***					
Between blocks	5	41·7	4·13	1·50	4·12
Between treatments	5	66·2	6·55	3·33	9·14
Interaction	25	10·1	—	0·365	—

Table 12. *The values of t, in the comparison of means, as calculated from the untransformed and the transformed data of Experiment III on* Leptinotarsa decemlineata

Comparison	Means for untransformed data		*t* without transformation	*t* from $(x+\frac{1}{2})^{\frac{1}{2}}$	*t* from $k^{-\frac{1}{2}} \sinh^{-1}(kx)^{\frac{1}{2}}$
$x_{1..}-x_{2..}$	362	30	+9·27**	+11·33**	+ 9·92**
$x_{1..}-x_{3..}$	362	224	+3·84**	+ 3·52**	+ 2·11*
$x_{1..}-x_{4..}$	362	11	+9·82**	+12·89**	+12·46**
$x_{2..}-x_{3..}$	30	224	−5·43**	− 7·81**	− 7·81**
$x_{2..}-x_{4..}$	30	11	+0·54	+ 1·56	+ 2·54*
$x_{3..}-x_{4..}$	224	11	+5·98**	+ 9·37**	+10·35**

treatment may be judged, there are shown, also in Table 12, the means for the untransformed data. The values of t falling beyond the 0·01 level of significance have been marked with two asterisks and the values beyond the 0·05 level with one. It can be seen that the transformation resulted in a profound alteration in the conclusions. Apparently on account of the dependence of variance on mean in untransformed data, the pooled estimate of variance was originally too low for the treatments which resulted in high populations and too high for the treatments which resulted in low populations. Thus, in the comparison of the first and third treatments, which appeared to have the two highest surviving populations, the value of t calculated from untransformed values was high. In the other extreme case, the comparison between the second and fourth treatments, the value of t, as calculated from untransformed values was very low. It can be seen further, that the first transformation only secured in part the modification in the value of t that was secured by the second transformation.

8. THE PROCEDURE OF TRANSFORMATION IN PRACTICE

The methods which were found applicable in the preceding discussion will now be illustrated in the transformation of the data shown in Table 7 (Experiment VII) on *Phlegethontius quinquemaculata*, of the same type as the experiments previously discussed in the present paper. The steps in the analysis will be set out with the purpose of providing a model for procedure in estimating the constant, k, which will be used to effect a transformation of the data so that the analysis of variance may be made.

Supposing that the experiment has been laid out with a repetition of each treatment in each block, the procedure of estimating k makes it first necessary to find the sum and the absolute difference of each pair of plots subjected to a given treatment in a given plot and then to sum the sums, $\sum_{i=1}^{n}\sum_{j=1}^{N}(x_{ij1}+x_{ij2})$, and the sums squared, $\sum_{i=1}^{n}\sum_{j=1}^{N}(x_{ij1}+x_{ij2})^2$, and also the differences squared, $\sum_{i=1}^{n}\sum_{j=1}^{N}(x_{ij1}-x_{ij2})^2$, over all such pairs and by substituting the results in (14) to find k. In the case being used for an illustration the two plots subjected to the first treatment in each block gave respectively 10 and 7, 20 and 14, 14 and 12, 10 and 23, 17 and 20, 14 and 13, so that

$$\sum_{j=1}^{N}(x_{1j1}+x_{1j2}) = (10+7)+(20+14)+(14+12)+\ldots = 174.$$

Similarly,

$$\sum_{j=1}^{N}(x_{1j1}+x_{1j2})^2 = (10+7)^2+(20+14)^2+\ldots = 5308$$

and similarly,

$$\sum_{j=1}^{N}(x_{1j1}-x_{1j2})^2 = (10-7)^2+(20-14)^2+\ldots = 228.$$

Of course, in estimating k the summations are not limited to one treatment but must be extended over all in the experiments. If this is done we find

$$\sum_{i=1}^{n}\sum_{j=1}^{N}(x_{ij1}+x_{ij2}) = 684, \quad \sum_{i=1}^{n}\sum_{j=1}^{N}(x_{ij1}+x_{ij2})^2 = 19{,}656, \quad \sum_{i=1}^{n}\sum_{j=1}^{N}(x_{ij1}-x_{ij2})^2 = 708.$$

From (14) we estimate $$k = \frac{2(708-684)}{19{,}656} = 0{\cdot}002,$$

and referring to Table 8, p. 250, use $k = 0{\cdot}00$ as the nearest value occurring there. Of course, in this case, the transformation is simply $x^{\frac{1}{2}}$.

Now from the above result it will be possible to replace the observed values of Table 7 with the corresponding transformed values from the first column of Table 8. Thus in Table 7 replace in the first row: 10, 20, 14, 10, 17 and 14, by 3·16, 4·47, 3·74, 3·16, 4·12 and 3·74. With such transformed values we can now proceed to carry out a routine analysis of variance which will be facilitated by working with the sum for each pair of plots in a given block with a given treatment. For example, the final analysis of variance for Experiment VII would be carried out with the values of Table 13.

Table 13. *Transformed and summed values to be used in the analysis of variance for Experiment VII on* P. quinquemaculata

Treatment	Block					
	1	2	3	4	5	6
1	5·81	8·21	7·20	7·96	8·59	7·35
2	7·44	7·90	7·74	8·24	8·94	6·26
3	1·00	4·06	2·73	2·41	1·73	3·00
4	3·97	5·91	3·73	4·48	4·48	3·41
5	3·97	3·97	4·18	2·00	3·14	4·45
6	6·32	8·56	7·87	6·77	10·20	8·51

9. Summary and conclusions

The foregoing work is a study of experimental results from seven field experiments on the control of insects. In such data, the standard deviation of the number of insects per plot varies with the mean. By the transformation, $x' = k^{-1}\sinh^{-1}(kx)^{\frac{1}{2}}$, where k is a constant and x an observation, the data were put in a form for which the standard deviation approached a constant independent of the mean. The estimation of the one constant, k, necessary for the transformation was made possible by the design of the experiments with repetition of treatments within blocks. In practice, the transformation gave good results so that

analysis of variance could be made. From the analysis of the transformed data, the results were found to differ markedly from those which would have been obtained from the untransformed data.

ACKNOWLEDGEMENTS

The present work was suggested to the writer by Prof. E. S. Pearson to whom the writer is further indebted for advice as the work progressed. The writer is further beholden to Dr R. W. B. Jackson, to Dr B. L. Welch and to Dr M. S. Bartlett for criticism and advice. Dr G. M. Stirrett and Mr F. A. Stinson very kindly permitted the use of unpublished data from their experiments carried out at Delhi, Ontario, on *Phlegethontius quinquemaculata* Haw.

APPENDIX

As has been said, the transformation of (10) was suggested by the method used by Tippett (1934, p. 61). The procedure is as follows.

It is required to find $x' = f(x)$, such that the standard deviation, $\sigma_{x'}$ of x', shall be approximately constant. Let us write

$$x' = f(M) + f'(M)(x - M) + \ldots, \tag{19}$$

where M is the expectation of x and whence, approximately,

$$(x' - M') = f'(M)(x - M), \tag{20}$$

where M' is the expectation of x'. Hence

$$\sigma_{x'}^2 = \{f'(M)\}^2 \sigma^2, \tag{21}$$

where σ is the standard deviation of the observations, x. Replacing $\sigma_{x'}$ in (21) by a constant, c, as is the purpose of our operation, and substituting for σ from equation (6), p. 247, we have

$$f'(M) = c(M + kM^2)^{-\frac{1}{2}}, \tag{22}$$

where k is, as has been previously discussed, a constant peculiar to our data. Integrating in (22),

$$f(M) = 2ck^{-\frac{1}{2}} \sinh^{-1}(kM)^{\frac{1}{2}}. \tag{23}$$

From (23) the form of the function suggested is $\sinh^{-1}(kx)^{\frac{1}{2}}$, but it is wise instead to use $k^{-\frac{1}{2}} \sinh^{-1}(kx)^{\frac{1}{2}}$, since the transformation then becomes identical, as shown in (11), with the established transformation, $x^{\frac{1}{2}}$, when $k = 0$.

As Tippett (1934) says: 'This derivation is not mathematically sound, and the result is only justified if on application it is found to be satisfactory.' The writer would have hesitated to have used it had it not already led to useful transformations in cases analogous to the present, namely to $x^{\frac{1}{2}}$ where x comes from a Poisson distribution, to $\sin^{-1} p^{\frac{1}{2}}$ where p comes from a binomial distribution and, according to Tippett, to $\tanh^{-1} r$, where r is the correlation coefficient.

REFERENCES

BARTLETT, M. S. (1936*a*). Square root transformation in analysis of variance. *J. Roy. Statist. Soc.* Suppl. **3**, 68–78.

—— (1936*b*). Some notes on insecticide tests in the laboratory and in the field. *J. Roy. Statist. Soc.* Suppl. **3**, 185–94.

BEALL, G. (1935). Study of arthropod populations by the method of sweeping. *Ecology*, **16**, 216–25.

—— (1939). Methods of estimating the population of insects in a field. *Biometrika*, **30**, 422–39.

—— (1940). The transformation of data from entomological field experiments. *Canadian Ent.* **72**, 168.

BEALL, G., STIRRETT, G. M. & CONNERS, I. L. (1939). A field experiment on the control of the European corn borer, *Pyrausta nubilalis* Hubn., by *Beauveria Bassiana* Vuill. II. *Sci. Agric.* **19**, 531–4.

BECKER, G. F. & VAN ORSTRAND, C. E. (1931). Hyperbolic functions (4th reprint). *Smithsonian Mathematical Tables.*

CLAPHAM, A. R. (1936). Over-dispersion in grassland communities and the use of statistical methods in plant ecology. *J. Ecol.* **24**, 232–51.

GEARY, R. C. (1935). The ratio of the mean deviation to the standard deviation as a test of normality. *Biometrika*, **27**, 310–32.

NAYER, P. P. N. (1936). An investigation into the application of Neyman and Pearson's L_1 test, with tables of percentage limits. *Statist. Res. Mem.* **1**, 38–51.

SNEDECOR, G. W. (1934). *Calculation and Interpretation of Analysis of Variance and Covariance.* Pp. 96. Ames, Iowa: Collegiate Press Inc.

'STUDENT' (1919). An explanation of deviations from Poisson's law in practice. *Biometrika*, **12**, 211–15.

TIPPETT, L. H. C. (1934). Statistical methods in textile research. Part 2. Uses of the binomial and Poisson distributions. *Shirley Inst. Mem.* **13**, 35–72.

WILLIAMS, C. B. (1937). The use of logarithms in the interpretation of certain entomological problems. *Ann. Appl. Biol.* **24**, 404–14.

INTERPOLATION FOR FRESH PROBABILITY LEVELS BETWEEN THE STANDARD TABLE LEVELS OF A FUNCTION

By J. B. SIMAIKA

1. The problem

A number of tables of probability functions exist, and more will no doubt before long be available, giving values for a variable x corresponding to a limited number of simple probability levels α. How far is it possible to obtain x rapidly for intermediate values of α?

The variable may be put into standardized form as the ratio of the deviation from the mean to the standard deviation; these two quantities (i.e. the mean and standard deviation) are often easy to obtain, whereas the probability integral may require extensive computation. Denote by u_α the standardized variable so that

$$u_\alpha = \frac{x_\alpha - \text{mean } x}{\text{standard deviation of } x}. \tag{1}$$

The question we shall consider is this: having full and accurate tables relating u_α and α for a standardized normal variable, denoted by U_α, can we use these values as auxiliary in obtaining u_α for any other function tabled only at a few probability levels and, having found an interpolating formula, what is its accuracy? In examining this point we shall compare the accuracy of the method with that of some other methods of deriving intermediate values of u_α.

2. General approach

It will be useful to consider first how far a general theoretical approach will take us. Let the variable x follow a probability law defined only by its cumulants $\kappa_r\ (r = 1, 2, \ldots)$; then the first four cumulants of the variable u become $0, 1, \gamma_1, \gamma_2$. It is known that the relation between u_α and α may be written symbolically

$$\alpha = \int_{-\infty}^{u_\alpha} \frac{1}{\sqrt{(2\pi)}} \exp\left[-\tfrac{1}{6}\gamma_1 \frac{d^3}{dx^3} + \tfrac{1}{24}\gamma_2 \frac{d^4}{dx^4} + \ldots\right] e^{-\frac{1}{2}x^2}\, dx, \tag{2}$$

while the same relation for a normal variable is

$$\alpha = \int_{-\infty}^{U_\alpha} \frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}x^2}\, dx. \tag{3}$$

Using equations (2) and (3) it has been shown by Cornish & Fisher (1937) that u_α can be approximated to by a parabolic curve in U_α and vice versa. If we

assume that κ_r for $r \geqslant 5$ is negligible, this expression, using a fifth degree parabola, is

$$u_\alpha = A + BU_\alpha + CU_\alpha^2 + DU_\alpha^3 + EU_\alpha^4 + FU_\alpha^5, \tag{4}$$

where

$$\begin{aligned}
A &= -\tfrac{1}{6}\gamma_1 - \tfrac{1}{12}\gamma_1\gamma_2 + \tfrac{17}{324}\gamma_1^3, \\
B &= 1 - \tfrac{1}{8}\gamma_2 + \tfrac{5}{36}\gamma_1^2 - \tfrac{29}{381}\gamma_2^2 - \tfrac{107}{288}\gamma_1^2\gamma_2 - \tfrac{1511}{7776}\gamma_1^4, \\
C &= \tfrac{1}{6}\gamma_1 + \tfrac{5}{24}\gamma_1\gamma_2 - \tfrac{53}{324}\gamma_1^3, \\
D &= \tfrac{1}{24}\gamma_2 - \tfrac{1}{18}\gamma_1^2 + \tfrac{1}{16}\gamma_2^2 - \tfrac{103}{288}\gamma_1^2\gamma_2 + \tfrac{211}{972}\gamma_1^4, \\
E &= -\tfrac{1}{24}\gamma_1\gamma_2 + \tfrac{1}{27}\gamma_1^3, \\
F &= -\tfrac{1}{128}\gamma_2^2 + \tfrac{7}{144}\gamma_1^2\gamma_2 - \tfrac{7}{216}\gamma_1^4.
\end{aligned}$$

Now as γ_1 and γ_2 tend to zero, B tends to unity and all other coefficients tend to zero, i.e. the curve of u_α as a function of U_α tends to the diagonal line

$$u_\alpha = U_\alpha.$$

Figs. 1–4 give these curves for different numbers of degrees of freedom for the commonly used statistical variables χ^2, χ, t and v (a transformation of z referred to below), expressed in standardized form.

With regard to the coefficients in (4) it may be remarked that large values of γ_2 do not increase them as much as large values of γ_1. Furthermore, when γ_1 is zero, the coefficients A, C and E vanish and the expression (4) becomes

$$u_\alpha = BU_\alpha + DU_\alpha^3 + FU_\alpha^5,$$

or

$$\frac{u_\alpha}{U_\alpha} = B + DU_\alpha^2 + F\{U_\alpha^2\}^2. \tag{5}$$

These broad results suggest that a good method of interpolation, when both γ_1 and γ_2 exist, is a Lagrange formula through the points

$$(u_{\alpha_i}, U_{\alpha_i}) \quad (i = 1, 2, \ldots).$$

Remembering that $u_{\alpha_i} = (x_{\alpha_i} - \text{mean } x)/(\text{S.D. of } x)$, from the practical point of view the interpolation can be carried out more expeditiously and without loss of accuracy by using a Lagrangian formula through

$$(x_{\alpha_i}, U_{\alpha_i}) \quad (i = 1, 2, \ldots).$$

When, however, γ_1 is zero as in the case of the t-distribution it would be better to take the Lagrange formula through the points

$$\left(\frac{u_{\alpha_i}}{U_{\alpha_i}}, U_{\alpha_i}^2\right) \text{ or alternatively through } \left(\frac{x_{\alpha_i}}{U_{\alpha_i}}, U_{\alpha_i}^2\right) \quad (i = 1, 2, \ldots),$$

the mean x being zero.

The accuracy of the method is likely to depend on the value of γ_1 and γ_2. For example, as seen in Fig. 1, linear interpolation between, say, $U_{0\cdot05}$ and $U_{0\cdot02}$ will be more accurate with $\nu = 18$ than $\nu = 3$, the γ's being smaller in the former case. Again, the curves will be more nearly linear if we take as variable

$$u = (\chi - \text{mean } \chi)/\sigma_\chi \quad \text{rather than} \quad u = (\chi^2 - \text{mean } \chi^2)/\sigma_{\chi^2}$$

because the γ's for the former are the smaller.

For the practical worker it will often be sufficient to use linear interpolation, i.e. to make use of two tabled probability levels only. For more accurate work

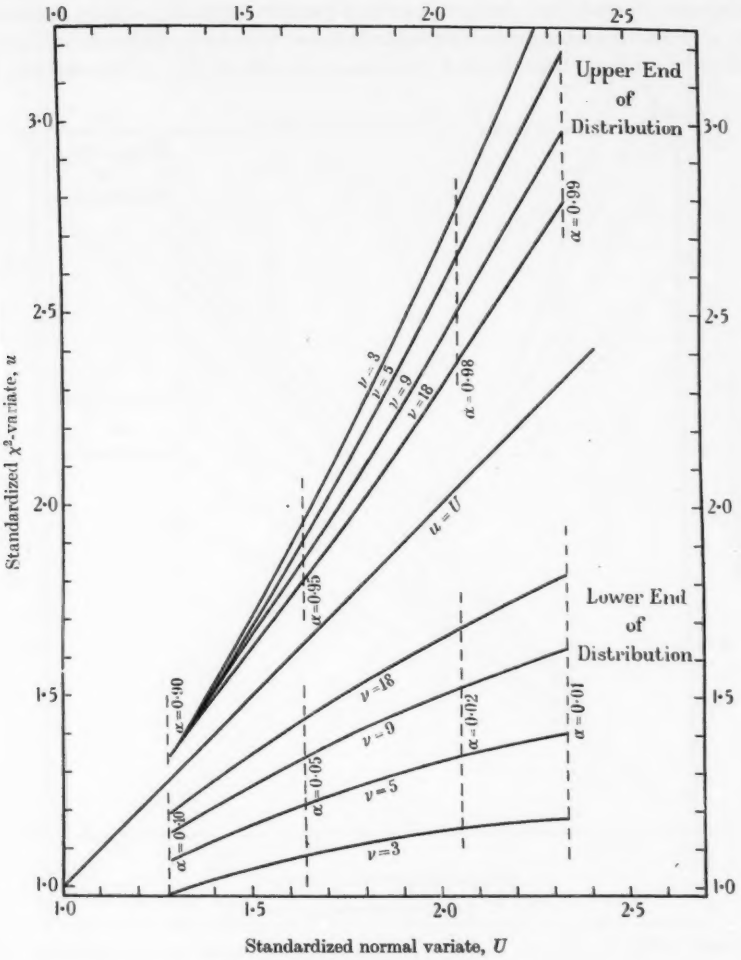


Fig. 1. Relation between $u(\chi^2)$ and U.

three or more levels can be used, but an increase beyond this is not in fact likely to lead to a gain in accuracy which will be worth the labour.

3. Comparison of methods

To compare the accuracy of the interpolation based on the polynomial expansion (4) with that obtained by other possible methods, we have considered the χ^2, t and v (beta)-probability distributions. For each, different methods of interpolation have been devised. In some a transformation of the variable has

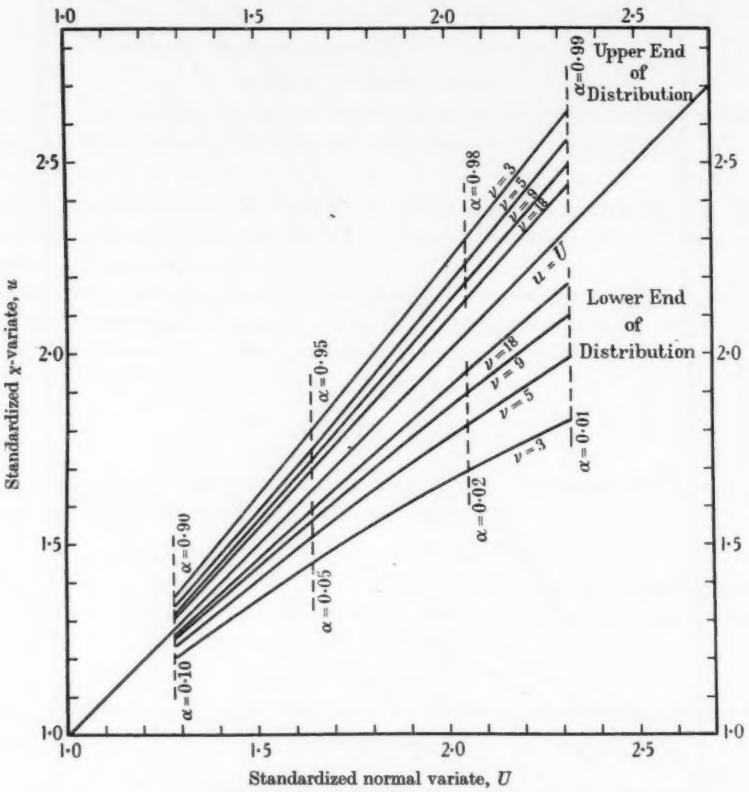


Fig. 2. Relation between $u(\chi)$ and U.

been used, while in others a transformation of the argument or a completely new argument has been considered.

The accuracy of each method in the range covered by the values $\alpha \leqslant 0{\cdot}10$ and $\alpha \geqslant 0{\cdot}90$ has been tested in the following way: between any two consecutive probability levels a number—not less than three—of intermediate values have been interpolated. These values were chosen to be those which could be obtained accurately from some other table. The greatest deviations in each interval are

given in Tables 1–3. The methods have been arranged according to the degree of accuracy obtained.

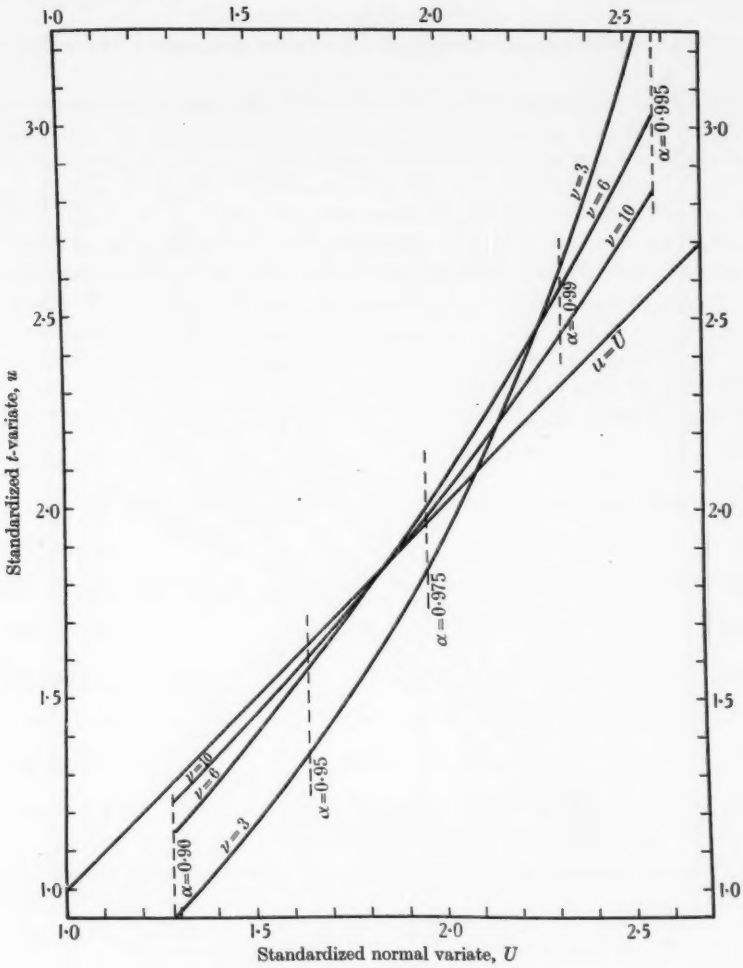


Fig. 3. Relation between $u(t)$ and U.

It remains to point out that, if the n tabulated probability levels are denoted by

$$\alpha_1 < \alpha_2 < \ldots < \alpha_n \leqslant 0{\cdot}10,$$

and if the interpolation is to be carried by using a quadratic or higher expression in the argument, it has been found that the interpolated value is always more

accurate when the probability levels used include as many as possible of the probability levels below α_x. Similarly, if

$$0{\cdot}90 \leqslant \alpha_1 < \alpha_2 < \ldots < \alpha_n,$$

it is better to use probability levels including as many as possible of those above α_x.

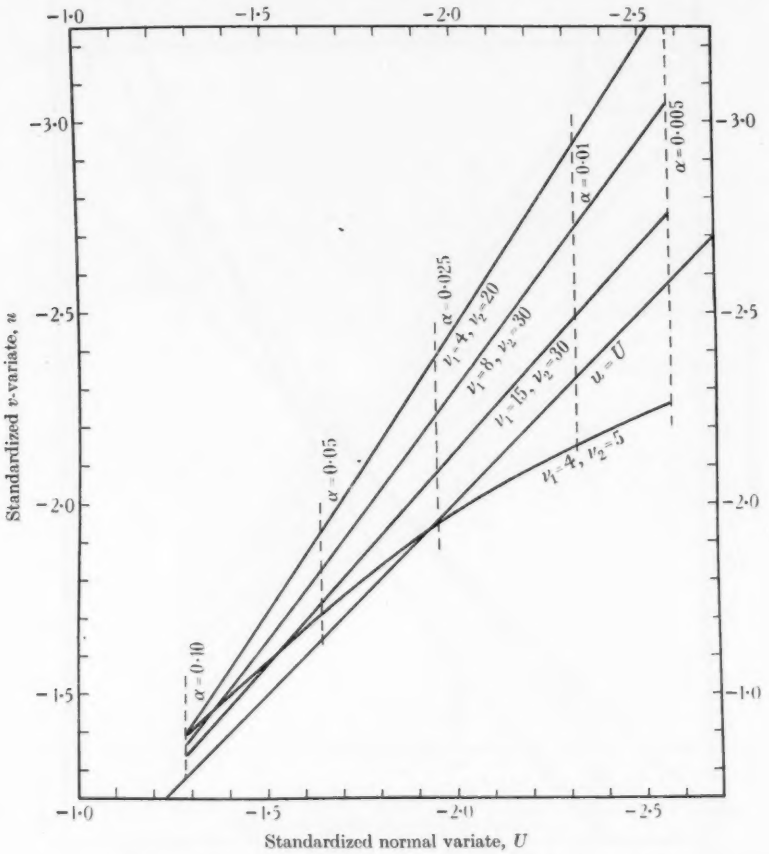


Fig. 4. Relation between $u(r)$ and U, where $p(v) = Cv^{\frac{1}{2}\nu_2 - 1}(1-v)^{\frac{1}{2}\nu_1 - 1}$.

4. The χ^2-probability function

The standardized form of χ^2 used in Fig. 1 is

$$u(\chi^2) = \frac{\chi^2 - \nu}{\sqrt{(2\nu)}}, \tag{6}$$

ν being the number of degrees of freedom. The γ_1 and γ_2 of this distribution are

$\sqrt{(8/\nu)}$ and $12/\nu$ respectively. If we consider χ itself instead of χ^2 we find that its standardized form has the approximate value, used in Fig. 2,

$$u(\chi) = \frac{\chi\sqrt{2} - \sqrt{(2\nu - 1)}}{1 - \frac{1}{8}\nu^{-1}}, \tag{7}$$

and its γ_1 and γ_2 are $(2\nu)^{-\frac{1}{2}} + O(\nu^{-2})$ and $O(\nu^{-2})$ respectively. Both these last quantities are smaller than those for χ^2, which suggests that interpolation will be more accurate using χ rather than χ^2. This can also be seen from Figs. 1 and 2.

The χ^2 probability levels have been tabulated by Fisher (1941) for $\alpha = 0{\cdot}01$, $0{\cdot}02$, $0{\cdot}05$, $0{\cdot}10$, ..., $0{\cdot}90$, $0{\cdot}95$, $0{\cdot}98$, $0{\cdot}99$ and are given to three decimal places. These levels were used in the interpolation.* The accuracy of the interpolated values obtained by the eight methods detailed below was checked either from the *Tables of the Incomplete Gamma Function* (Karl Pearson, 1922) or from *Tables or Statisticians and Biometricians*, Part I, Table XII (Karl Pearson, 1930).

The greatest deviations, $\delta_m (m = 1, 2, \ldots, 8)$, obtained using in all eight different methods, are given for $\nu = 3, 5, 9$ and 18 and for intervals of α: $(0{\cdot}01, 0{\cdot}02)$, $(0{\cdot}02, 0{\cdot}05)$, $(0{\cdot}05, 0{\cdot}10)$, $(0{\cdot}90, 0{\cdot}95)$, $(0{\cdot}95, 0{\cdot}98)$ and $(0{\cdot}98, 0{\cdot}99)$ in Table 1.

Method 1.

$$\chi^2 = A + B\log\alpha \quad (\alpha < 0{\cdot}10), \qquad \chi^2 = A + B\log(1-\alpha) \quad (\alpha > 0{\cdot}90).$$

Method 2.

$$u(\chi) = A + B\log\alpha \quad (\alpha < 0{\cdot}10), \qquad u(\chi) = A + B\log(1-\alpha) \quad (\alpha > 0{\cdot}90).$$

Method 3. $$u(\chi^2) = A + BU.$$

Method 4.

$$u(\chi^2) - U = A + B\log\alpha \ (\alpha < 0{\cdot}10), \qquad u(\chi^2) - U = A + B\log(1-\alpha) \ (\alpha > 0{\cdot}90).$$

Method 5. $$u(\chi) = A + BU.$$

Method 6.
$$\chi^2 = A + B\log\alpha + C\log^2\alpha \quad (\alpha < 0{\cdot}10),$$
$$\chi^2 = A + B\log(1-\alpha) + C\log^2(1-\alpha) \quad (\alpha > 0{\cdot}90).$$

Method 7. $$u(\chi^2) = A + BU + CU^2.$$

Method 8. $$u(\chi) = A + BU + CU^2.$$

The best linear interpolation is that provided by method 5 and the best quadratic one is given by method 8. Both these methods are those suggested by the general approach.

* Certain additional levels are given in a recently published table computed by Catherine M. Thompson (1941).

Table 1. *Greatest deviations in the interpolation of χ^2 and χ.*

ν	γ_1	γ_2	Interval	α	True χ^2	$\lvert\delta_1\rvert$	$\lvert\delta_2\rvert$	$\lvert\delta_3\rvert$	$\lvert\delta_4\rvert$	$\lvert\delta_5\rvert$	$\lvert\delta_6\rvert$	$\lvert\delta_7\rvert$	$\lvert\delta_8\rvert$
3	1·63	4·00	(0·02, 0·05)	0·02 997	0·245	0·006	0·007	0·011	0·012	0·004	0·003	0·003	0·001
			(0·05, 0·10)	0·07 890	0·490	15	8	10	1	4	4	3	0
			(0·90, 0·95)	0·92 810	7·000	4	3	27	14	5	1	0	1
			(0·95, 0·98)	0·97 071	9·000	5	33	36	10	6	0	0	2
			(0·98, 0·99)	0·98 827	11·000	1	13	34	5	1	2	1	2
5	1·26	2·40	(0·01, 0·02)	0·01 353	0·632	8	5	6	3	2	6	0	0
			(0·02, 0·05)	0·03 340	0·949	23	13	15	3	5	5	1	0
			(0·05, 0·10)	0·06 150	1·265	19	11	11	0	3	4	0	1
			(0·90, 0·95)	0·92 808	10·119	11	32	26	12	5	3	1	1
			(0·95, 0·98)	0·96 544	12·017	13	39	32	12	6	2	1	0
			(0·98, 0·99)	0·98 579	14·230	5	18	17	6	4	2	1	1
9	0·94	1·33	(0·01, 0·02)	0·01 060	2·121	4	2	2	2	0	1	1	0
			(0·02, 0·05)	0·03 452	2·970	34	21	19	0	6	6	0	1
			(0·05, 0·10)	0·07 705	3·818	33	24	17	6	5	8	1	0
			(0·90, 0·95)	0·93 312	16·000	19	38	17	11	5	4	0	0
			(0·95, 0·98)	0·96 483	18·000	22	46	31	12	7	4	1	1
			(0·98, 0·99)	0·98 735	21·000	8	18	14	5	4	2	1	0
18	0·67	0·67	(0·01, 0·02)	0·01 167	7·200	14	9	5	2	1	4	1	1
			(0·02, 0·05)	0·02 793	8·400	47	32	20	2	5	7	0	1
			(0·05, 0·10)	0·07 482	10·200	48	34	18	5	6	11	1	0
			(0·90, 0·95)	0·93 159	27·600	34	52	25	10	5	7	1	1
			(0·95, 0·98)	0·96 255	30·000	34	55	28	10	6	6	1	1
			(0·98, 0·99)	0·98 589	33·600	15	26	13	5	4	3	1	1

For methods associated with subscripts to δ, see p. 269.

Table 2. *Greatest deviations in the interpolation of t*

ν	γ_2	Interval	α	True t	$\lvert\delta_1\rvert$	$\lvert\delta_2\rvert$	$\lvert\delta_3\rvert$	$\lvert\delta_4\rvert$	$\lvert\delta_5\rvert$	$\lvert\delta_6\rvert$	$\lvert\delta_7\rvert$	$\lvert\delta_8\rvert$
3	∞	(0·005, 0·01)	0·00 692	5·196	0·039	0·040	0·035	0·020	0·011	0·008	0·007	0·004
		(0·01, 0·025)	0·01 430	3·969	64	52	42	21	15	7	6	2
		(0·025, 0·05)	0·03 261	2·848	28	24	16	7	8	5	3	0
		(0·050, 0·10)	0·07 285	1·954	25	23	11	4	7	3	3	1
6	3·00	(0·005, 0·01)	0·00 714	3·413	11	8	5	3	2	1	1	1
		(0·01, 0·025)	0·01 517	2·820	18	14	6	3	2	1	1	1
		(0·025, 0·05)	0·03 874	2·128	8	6	1	1	2	2	0	1
		(0·050, 0·10)	0·06 820	1·719	9	8	1	1	1	1	0	0
10	1·00	(0·005, 0·01)	0·00 625	3·038	5	4	1	2	1	1	1	1
		(0·01, 0·025)	0·01 639	2·476	8	11	1	1	1	1	1	1
		(0·025, 0·05)	0·03 844	1·972	4	3	2	1	1	1	0	0
		(0·050, 0·10)	0·07 246	1·581	4	4	5	0	1	1	0	0

For methods associated with subscripts to δ, see p. 271.

5. THE t-PROBABILITY FUNCTION

The standardized form of the t-probability function used in Fig. 3 is

$$u(t) = t\sqrt{\left(\frac{\nu-2}{\nu}\right)}, \tag{8}$$

ν being the number of degrees of freedom. The values of γ_1 and γ_2 are zero and $6/(\nu-4)$ respectively.

Percentage levels for t have been tabulated by Fisher (1941) for $\alpha = 0{\cdot}005$, $0{\cdot}01$, $0{\cdot}025$, $0{\cdot}05$, $0{\cdot}10$, ... and are given to three decimal places. The level here defined as α is half the figure given in Fisher's table, i.e.

$$\alpha = \int_{-\infty}^{t_\alpha} p(t)\,dt. \tag{9}$$

The values of t used in checking those obtained by interpolation are taken from *Tables of the Incomplete Beta-Function* (Karl Pearson, 1934), where

$$\left.\begin{aligned} &I_x(p,q) = \alpha, \\ &p = \tfrac{1}{2}\nu, \quad q = \tfrac{1}{2} \quad \text{and} \quad x = \frac{1}{1+t^2/\nu} \end{aligned}\right\}, \tag{10}$$

ν being the number of degrees of freedom.

The greatest deviations found in the following eight methods are denoted by δ_m ($m = 1, 2, \ldots, 8$), and are given in Table 2 for $\nu = 3, 6$ and 10 and for the intervals $(0{\cdot}005, 0{\cdot}01)$, $(0{\cdot}01, 0{\cdot}025)$, $(0{\cdot}025, 0{\cdot}05)$, $(0{\cdot}05, 0{\cdot}10)$.

Method 1. $u(t) = A + BU.$

Method 2. $u(t) - U = A + B\log\alpha.$

Method 3. $t = A + B\log\alpha.$

Method 4. $u(t) = AU + BU^3.$

Method 5. $u(t) = A + BU + CU^2.$

Method 6. $t = A + B\log\alpha + C\log^2\alpha.$

Method 7. $u(t) - U = A + B\log\alpha + C\log^2\alpha.$

Method 8. $u(t) = AU + BU^3 + CU^5.$

From the methods that make use of only two probability levels, method 4 is the one that gives the highest accuracy and from those that make use of three probability levels, method 8 is the best. These two methods are those suggested by the general approach.

6. THE v OR BETA-DISTRIBUTION

This variable, which is related to R. A. Fisher's z, is defined as

$$v = \frac{S_2}{S_1+S_2} = \frac{\nu_2}{\nu_2+\nu_1 e^{2z}}, \tag{11}$$

where S_1 is a sum of squares of normal variates based on ν_1 degrees of freedom and S_2 an independent sum of squares based on ν_2 degrees of freedom. The elementary probability law of v is

$$p(v) = \{B(\tfrac{1}{2}\nu_1, \tfrac{1}{2}\nu_2)\}^{-1} v^{\frac{1}{2}\nu_2 - 1}(1-v)^{\frac{1}{2}\nu_1 - 1}. \tag{12}$$

Table 3. *Greatest deviations in the interpolation of v*

ν_1	ν_2	γ_1	γ_2	Interval	α	True v	$\lvert\delta_1\rvert$	$\lvert\delta_2\rvert$	$\lvert\delta_3\rvert$	$\lvert\delta_4\rvert$	$\lvert\delta_5\rvert$	$\lvert\delta_6\rvert$	$\lvert\delta_7\rvert$
4	20	−0·92	+0·79	(0·005, 0·01)	0·00 838	0·52000	0·00043	0·00016	0·00006	0·00007	0·00006	0·00005	0·00008
				(0·01, 0·05)	0·02 607	0·59000	296	76	34	35	7	27	1
				(0·05, 0·10)	0·06 901	0·66000	79	5	28	14	6	10	1
8	30	−0·61	+0·25	(0·005, 0·01)	0·00 769	0·53000	25	4	7	16	9	5	8
				(0·01, 0·05)	0·02 711	0·59000	228	36	24	35	4	13	1
				(0·05, 0·10)	0·06 645	0·64000	67	4	16	16	2	5	2
15	30	−0·28	−0·12	(0·005, 0·01)	0·00 666	0·41000	31	3	3	8	9	7	8
				(0·01, 0·05)	0·02 826	0·47000	240	40	22	42	3	9	3
				(0·05, 0·10)	0·07 463	0·52000	78	9	1	20	1	9	1
8	15	−0·33	−0·26	(0·005, 0·01)	0·00677	0·30000	53	19	20	13	3	3	3
				(0·01, 0·05)	0·02 519	0·37000	418	128	114	67	11	22	2
				(0·05, 0·10)	0·06 972	0·44000	126	27	16	29	2	15	3
4	5	−0·16	−0·77	(0·005, 0·01)	0·00 796	0·09000	80	32	58	29	16	5	5
				(0·01, 0·05)	0·02 723	0·15000	838	375	534	181	93	19	12
				(0·05, 0·10)	0·07 421	0·23000	278	123	144	78	33	8	11

For methods associated with subscripts to δ, see p. 273.

Low values of v correspond to high values of z. The standardized form of v used in Fig. 4 is

$$u(v) = \frac{v - \dfrac{\nu_2}{\nu_1+\nu_2}}{\dfrac{1}{\nu_1+\nu_2}\sqrt{\left(\dfrac{2\nu_1\nu_2}{\nu_1+\nu_2+2}\right)}}, \tag{13}$$

and the values of γ_1 and γ_2 are

$$\left.\begin{aligned} \gamma_1 &= \frac{\nu_1-\nu_2}{\nu_1+\nu_2+4}\sqrt{\left(\frac{8(\nu_1+\nu_2+2)}{\nu_1\nu_2}\right)}, \\ \gamma_2 &= \frac{12\{\nu_1^2(\nu_1+2)+\nu_2^2(\nu_2+2)-2\nu_1\nu_2(\nu_1+\nu_2+4)\}}{\nu_1\nu_2(\nu_1+\nu_2+4)\,(\nu_1+\nu_2+6)}. \end{aligned}\right\} \tag{14}$$

The cases considered here are those most generally met in tests of significance with $\nu_1 < \nu_2$ and therefore $\gamma_1 < 0$. γ_2 is sometimes positive and sometimes negative.

Tables giving values of v for $\alpha = 0{\cdot}005$, 0·01, 0·025, 0·05, 0·10, 0·25 and 0·50 to five decimal places have recently been published (Catherine M. Thompson, 1941). The corresponding upper probability levels can be obtained by entering the tables with ν_1 and ν_2 transposed and taking $1-v$ for v.

The values of v used in checking the interpolation are taken from the *Tables of the Incomplete Beta-Function* (Karl Pearson, 1934).

The greatest deviations δ_m $(m = 1, 2, ..., 7)$, found in the following seven methods are given in Table 3 for the following pairs of values of ν_1 and ν_2: (4, 20) (8, 30) (15, 30) (8, 15) (4, 5) and for the intervals (0·005, 0·01), (0·01, 0·05), (0·05, 0·10) of α.

Method 1.

$$v = A + B\log\alpha \quad (\alpha < 0{\cdot}10), \qquad v = A + B\log(1-\alpha) \quad (\alpha > 0{\cdot}90).$$

Method 2.

$$u(v) - U = A + B\log\alpha \quad (\alpha < 0{\cdot}10), \qquad u(v) - U = A + B\log(1-\alpha) \quad (\alpha > 0{\cdot}90).$$

Method 3. $$u(v) = A + BU.$$

Method 4.
$$v = A + B\log\alpha + C\log^2\alpha \quad (\alpha < 0{\cdot}10),$$
$$v = A + B\log(1-\alpha) + C\log^2(1-\alpha) \quad (\alpha > 0{\cdot}90).$$

Method 5.
$$u(v) - U = A + B\log\alpha + C\log^2\alpha \quad (\alpha < 0{\cdot}10),$$
$$u(v) - U = A + B\log(1-\alpha) + C\log^2(1-\alpha) \quad (\alpha > 0{\cdot}90).$$

Method 6. $$u(v) = A + BU + CU^2.$$

Method 7. $$u(v) = A + BU + CU^2 + DU^3.$$

Here also the best linear method and the best quadratic one are those suggested by the general approach, namely, methods 3 and 6. Method 7, a cubic in U, was used only because the great accuracy of the tabulated v-function justified the computation. This cubic interpolation gives errors of the order of 0·0001 even for numbers of degrees of freedom as small as $\nu_1 = 4$, $\nu_2 = 5$. If the interval between $\alpha = 0{\cdot}01$ and $0{\cdot}05$ were broken into two parts at 0·025 (a level given in the Thompson tables) the corresponding errors δ_3 would be considerably reduced.

7. SOME NUMERICAL EXAMPLES, USING LINEAR INTERPOLATION

(*a*) *Interpolation for a χ^2 level.* Suppose that we calculate the upper $2\frac{1}{2}$ % level ($\alpha = 0{\cdot}975$) for χ^2 with $\nu = 5$ degrees of freedom, using tabled values of the upper 2 % and 5 % levels. We require the 2 %, $2\frac{1}{2}$ %, and 5 % levels of the standardized normal variable as well as the two χ^2 levels. The relevant data are shown in the following table; as pointed out above, there is no need to calculate the standardized form of either χ or χ^2.

α	U_α	χ^2_α	χ_α
0·98	2·0537	13·388	3·659
0·975	1·9600	?	?
0·95	1·6449	11·070	3·327

By linear interpolation, i.e. using method 3 of p. 269, we have

$$\chi^2_{0\cdot975} = 11\cdot070 + (13\cdot388 - 11\cdot070) \times \frac{1\cdot9600 - 1\cdot6449}{2\cdot0537 - 1\cdot6449}$$
$$= 12\cdot857.$$

Interpolating similarly for χ, i.e. using method 5 of p. 269, we find

$$\chi_{0\cdot975} = 3\cdot583 \quad \text{or} \quad \chi^2_{0\cdot975} = 12\cdot838.$$

The correct value taken from Miss Thompson's table (1941) is

$$\chi^2_{0\cdot975} = 12\cdot8325.$$

It is seen, as expected on theoretical grounds and as evidenced in the comparisons of Table 1, that method 5 is the more accurate of the two.

(*b*) *Interpolation for a t level.* Suppose that we calculate the value of t corresponding to $\alpha = 0\cdot0125$ (as defined in equation (9)), with $\nu = 6$, using the tabled levels for $\alpha = 0\cdot005$ and $0\cdot025$. The data required are shown below, the values of t being taken from the table on p. 300 of this issue.

α	U_α	t_α	t_α/U_α	U_α^2
0·005	−2·5758	−3·7074	1·4393	6·6347
0·0125	−2·2414	?	?	5·0239
0·025	−1·9600	−2·4469	1·2484	3·8416

Again, it is not necessary to calculate the standardized values of t_α, for even when using the relation of method 4, which assumes

$$\frac{t}{\sigma_t U} = A + BU^2, \tag{15}$$

the transference of σ_t to the right-hand side of the equation will only modify the constants A and B whose values are not directly determined in the interpolation process.

Using method 1 (t a linear function of U), it is found that

$$t_{0\cdot0125} = -3\cdot023.$$

Using method 4 (t/U a linear function of U^2), it is found that

$$t_{0\cdot0125} = -2\cdot979.$$

The correct value is $-2\cdot969$. As can be seen in Fig. 3, from the stretch of the curve for $\nu = 6$ between $\alpha = 0\cdot975$ and $0\cdot995$, the interval chosen is too long for satisfactory linear interpolation. The use of formula (15) improves matters, but is still not satisfactory. With Fisher's tables we could, of course, interpolate between the levels $\alpha = 0\cdot010$ and $0\cdot025$; doing this with method 4, it is found that

$$t_{0\cdot0125} = -2\cdot971$$

a distinctly better value.

(*c*) *Interpolation for a beta-distribution percentage level.* Take the case $\nu_1 = 8$, $\nu_2 = 30$ and suppose it is wished to find the $2\frac{1}{2}$ % level from a knowledge of the 1 % and 5 % levels. The data required are as follows, $v_{0\cdot01}$ and $v_{0\cdot05}$ being taken from Miss Thompson's tables.

α	U_α	v_α
0·01	2·3263	0·54170
0·025	1·9600	?
0·05	1·6449	0·62332

Then, using linear interpolation, i.e. method 3 of p. 273, we have

$$u_{0\cdot025} = 0\cdot62332 + (0\cdot62332 - 0\cdot54170) \times \frac{1\cdot9600 - 1\cdot6449}{2\cdot3263 - 1\cdot6449}$$
$$= 0\cdot58558.$$

The correct value taken from the same tables is 0·58582. It will be seen from Fig. 4 that the intervals $\alpha = 0\cdot005, 0\cdot010, 0\cdot025, 0\cdot050, 0\cdot100$ of Miss Thompson's tables are likely to form a satisfactory framework for subtabulation if this is needed.

8. SUBTABULATION OF EXISTING TABLES

These methods have been used to produce the following enlargements of existing tables; but it is not at present possible to arrange for their publication.

(*a*) *Table of χ^2 percentage levels.* Method 7 (p. 269) was used, as it is almost as accurate as method 8 and less laborious. The table calculated gives χ^2_α to 3 decimal places for $\nu = 1\,(1)\,30$, and for $\alpha = 0\cdot010\,(0\cdot002)\,0\cdot020; 0\cdot020\,(0\cdot005)\,0\cdot050; 0\cdot05\,(0\cdot01)\,0\cdot10$; and for the corresponding levels at the upper end of the distribution, i.e. for $\alpha' = 1-\alpha$.

(*b*) *Table of t percentage levels.* Method 8 (p. 271) was used for $\nu = 3, \ldots, 6$, and method 6 for $\nu = 7\,(1)\,30; 40, 60$ and 120. Exact levels were calculated for $\nu = 1, 2$. The tables were computed to 3 decimal places and for

$\alpha = 0\cdot005\,(0\cdot001)\,0\cdot010; 0\cdot0100\,(0\cdot0025)\,0\cdot0250; 0\cdot025\,(0\cdot005)\,0\cdot050; 0\cdot05\,(0\cdot01)\,0\cdot10.$

9. INTERPOLATION FOR THE PROBABILITY INTEGRAL α, GIVEN u_α.

As we have already mentioned, the variable U_α can be expressed as a polynomial in u_α. This suggests that interpolation for U_α, given u_α, is as easy and as accurate as the interpolation discussed above. Hence to interpolate for α, given the value of u_α and certain tabled levels $u_{\alpha_1}, u_{\alpha_2}, \ldots$, we first interpolate for U_α and then find the value of α from appropriate tables of the normal probability integral, e.g. *Tables for Statisticians and Biometricians*, Part I, Table II.

10. Conclusion

It has been shown how accurate values of the probability levels of a statistical variate may be interpolated between standard tabled values by using the standardized normal variate as auxiliary. For many purposes linear interpolation is adequate; for others a second order Lagrangian formula may be preferred. The accuracy of the result depends, of course, on the closeness of the actual probability law to the normal law and on the size of the intervals between the tabled levels. The method has been illustrated on examples from the χ^2, t and v (beta) distributions.

The method has been used to provide a subtabulation of existing tables of percentage points for χ^2 and t, but it is not possible to have these tables printed with the present contribution.

Finally I should like to express my thanks to Dr B. L. Welch of University College, London, for originally suggesting the problem to me.

REFERENCES

Cornish, E. A. & Fisher, R. A. (1937). Moments and cumulants in the specification of distributions. *Rev. Inst. Int. Statist.* no. 4.

Fisher, R. A. (1941). *Statistical Methods for Research Workers*, 8th Edition. Edinburgh: Oliver & Boyd.

Pearson, K. (1922). *Tables of the Incomplete Gamma-Function.*

—— (1930). *Tables for Statisticians and Biometricians*, Part I.

—— (1934). *Tables of the Complete and Incomplete Beta-Function.*

Thompson, Catherine M. (1941). *Biometrika*, **32**, 168–81

PARTIAL RANK CORRELATION

By M. G. KENDALL

1. In interpreting an observed dependence between two qualities we are constantly faced with the question whether an association (correlation) of A with B is really due to the associations (correlations) of each with a third quality C. This has led naturally to the theories of partial association and correlation, which attempt to decide the matter by the consideration of subpopulations in which the variation of C is eliminated. An analogous problem arises in ranking work but, so far as I know, has not previously been considered. For example, if a number of men are ranked according to mathematical and musical aptitude and there appears a significant rank correlation, it is natural to inquire whether this may be attributable to the correlation of both with some more fundamental quality such as intelligence. The object of this paper is to propose a coefficient of partial rank correlation which has a natural meaning and may be found useful for investigations requiring this type of decision.

2. As a preliminary it may be worth examining what can be done in this direction with the Spearman rank correlation coefficient ρ. If there are three rankings denoted by 1, 2 and 3, we may find the three coefficients ρ_{12}, ρ_{13}, ρ_{23}. It is tempting to apply to these coefficients the formulae of product moment partial correlation such as

$$\rho_{23.1} = \frac{\rho_{23} - \rho_{12}\rho_{13}}{(1-\rho_{12}^2)^{\frac{1}{2}}(1-\rho_{13}^2)^{\frac{1}{2}}}, \tag{1}$$

and to *define* $\rho_{23.1}$ as the partial rank correlation of 2 and 3 'when 1 is constant'. There is clearly very little justification for such a procedure, and it is far from easy to explain just what $\rho_{23.1}$ means. In fact, the only defence of formula (1) that can bear critical examination is, I think, that it is an approximation to a second possibility, as follows:

3. There can be no such thing as a rank correlation in a continuous population (the members of which are not even denumerable) but we can speak with genuine meaning of a grade correlation. A well-known result due to Karl Pearson states that in a normal bivariate population with correlation ρ_p the grade correlation ρ_g is given by

$$\rho_p = 2\sin\frac{\pi}{6}\rho_g. \tag{2}$$

The Spearman coefficient ρ may be regarded as a sample grade correlation. If, therefore, we take ρ as an estimate of ρ_g we may find ρ_p from (2). For three rankings we may then obtain the three values of ρ_p, apply the ordinary product moment partial formula, and so obtain a partial coefficient. Since x and $2\sin\frac{\pi x}{6}$ do not

differ by more than a small amount for $|x| \leqslant 1$, we might even apply formula (1) direct to the values of ρ without bothering to transform them into ρ_p by equation (2).

4. Such a procedure, again, is open to fairly obvious objections. Apart from the all-too-facile assumption of normality and the error involved in using Spearman's ρ from a small sample to estimate the grade correlation in a parent, the fact remains that we arrive, not at a partial rank coefficient, but at an estimate of a partial product moment coefficient in a normal population.

Perhaps there are cases where this is a reasonable objective based on reasonable assumptions but it is evidently unsatisfactory for general ranking purposes.

5. In a previous paper (1938) I defined an alternative coefficient of rank correlation τ which may be generalized to include the case when pairs of individuals are compared separately (Kendall & Babington Smith, 1940). It will be convenient for present purposes to redefine τ in a slightly different manner so that the results obtained below may again be immediately generalized to the case of paired comparisons. Consider the two rankings of six

$$\begin{array}{lllllll} 1: & 1 & 4 & 3 & 2 & 6 & 5 \\ 2: & 3 & 2 & 4 & 1 & 5 & 6. \end{array} \qquad (3)$$

There are $\binom{6}{2} = 15$ possible pairs of ranks in each ranking, viz. 12, 13, ..., 16, 23, 24, ..., 56. We write them down as in expression (4) below. Any order of the pairs will serve, and it is immaterial whether any pair is written as *ab* or *ba*; but for practical convenience they may be written in the natural order indicated in the last sentence but one. This arrangement I call the recorded order.

We then consider the occurrence of each pair in the ranking 1. If a pair occurs in that ranking in the order in which we have recorded it, we write a plus below the recorded order underneath the pair concerned; in the contrary case we write a minus. Ranking 1 of expression (3) will then give

Recorded order:	(12)	(13)	(14)	(15)	(16)	(23)	(24)	(25)	(26)	(34)	(35)	(36)	(45)	(46)	(56)
Ranking 1:	+	+	+	+	+	−	−	+	+	−	+	+	+	+	−

(4)

Here, for example, the pair (15) occurs in that order in ranking 1 and so is denoted by a +, whereas the pair (24) occurs as 42 and is denoted by a −.

Consider now ranking 2. The members of ranking 1 which are ranked 1, 2 correspond to members in ranking 2 ranked as 3, 1. This is in the reverse order to that of the pair 13 in the recorded order, so (starting a new row of signs corresponding to ranking 2) we write a minus under the recorded pair (12). Similarly, the pair in ranking 2 corresponding to 15 in ranking 1 is 36. This is in the same order as the recorded pair, so we write a plus below the existing plus under the recorded pair (15). The pair in ranking 2 corresponding to 23 in ranking 1 is 41. This is in the reverse order of the recorded pair, so we write a minus below the

recorded pair (23) in the row of signs corresponding to ranking 2. And so on. This takes rather a long time to explain but the process is really very simple. The array corresponding to expressions (3) is then

Recorded order:	(12)	(13)	(14)	(15)	(16)	(23)	(24)	(25)	(26)	(34)	(35)	(36)	(45)	(46)	(56)
Ranking 1:	+	+	+	+	+	−	−	+	+	−	+	+	+	+	−
Ranking 2:	−	+	−	+	+	−	−	+	+	+	+	+	+	+	+

(5)

Now in expression (5) there are eleven cases in which both rankings have the same sign and therefore $15-11=4$ in which they have the opposite sign. The coefficient τ is then given by

$$\tau_{12}=\frac{11-4}{15}=\frac{7}{15}.$$

Generally, if there are, in two rankings of n arrayed as above, S_1 cases of the same sign and S_2 of opposite sign

$$\tau=\frac{2(S_1-S_2)}{n(n-1)}$$

$$=\frac{4S_1}{n(n-1)}-1=1-\frac{4S_2}{n(n-1)}. \tag{6}$$

If we arrange ranking 1 in the natural order 1, ..., n, then every case in which there are the same signs in expression (5) corresponds to a case in which pairs in ranking 2 are in the natural order; and every case of different sign to one in which the pairs in ranking 2 are in the reverse of the natural order. The definition of (6) thus accords with the one originally given in my 1938 paper. It is often convenient to take one ranking to be the natural order 1, ..., n so that the first row of signs in (5) are all positive. For instance, on rearrangement of the rankings in (3) we have

1:	1	2	3	4	5	6
2:	3	1	4	2	6	5

(7)

and the array of paired comparisons becomes

Recorded order:	(12)	(13)	(14)	(15)	(16)	(23)	(24)	(25)	(26)	(34)	(35)	(36)	(45)	(46)	(56)
Ranking 1:	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Ranking 2:	−	+	−	+	+	+	+	+	+	−	+	+	+	+	−

(8)

Here again $S_1=11$, $S_2=4$, $\tau_{12}=\frac{7}{15}$.

6. Consider now three rankings, of which the first may be taken to be the natural order 1, ..., n, for example,

1:	1	2	3	4	5	6
2:	3	1	4	2	6	5
3:	4	2	1	6	3	5

(9)

The corresponding array of pairs is

Recorded order:	(12)	(13)	(14)	(15)	(16)	(23)	(24)	(25)	(26)	(34)	(35)	(36)	(45)	(46)	(56)
Ranking 1:	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Ranking 2:	−	+	−	+	+	+	+	+	+	−	+	+	+	+	−
Ranking 3:	−	−	+	−	+	−	+	+	+	+	+	+	−	−	+

(10)

For the coefficients τ we have

$$\tau_{12} \text{ (as above)} = \frac{7}{15},$$

$$\tau_{13} = \frac{9-6}{15} = \frac{3}{15},$$

$$\tau_{23} = \frac{7-8}{15} = -\frac{1}{15},$$

in the last case S_1 being as usual the number of cases in which pairs in rankings 2 and 3 have the same sign.

Consider now the fourfold table setting out the occurrences of + and − signs in the rows of expression (10) corresponding to rankings 2 and 3:

Ranking 3 \ Ranking 2	+	−	Total
+	6	5	11
−	3	1	4
Total	9	6	15

(11)

Here, for example, there are six cases in which pairs of ranks have the same (positive) sign, five in which ranking 2 is negative while ranking 3 is positive, and so on.

Generally, if for three rankings of n the table is of the form

Ranking 3 \ Ranking 2	+	−	Total
+	a	b	$a+b$
−	c	d	$c+d$
Total	$a+c$	$b+d$	N

(12)

I define the partial rank correlation of 2 and 3 on 1 as

$$r_{23.1} = \frac{ad - bc}{+\sqrt{\{(a+b)(c+d)(a+c)(b+d)\}}} \tag{13}$$

$$= \sqrt{\frac{\chi^2}{N}}, \tag{14}$$

where $N = \binom{n}{2}$ and χ^2 is the ordinary mean square contingency for the fourfold table.

In the particular case here considered the coefficient is

$$r_{23.1} = \frac{6 - 15}{\sqrt{(4 \,.\, 11 \,.\, 9 \,.\, 6)}} = -0{\cdot}185,$$

as compared with $r_{23} = -0{\cdot}067.$

In the case when ranking 1 is not the natural order 1, ..., n the same principles apply, but in considering the communalities of 2 and 3 we count as contributing to the a-cell in (12) the pairs which are themselves of the same sign and are also of the same sign as the term in ranking 1; and so on.

7. The partial r defined in equation (13) is a coefficient of association in the 2×2 table suggested by Yule (1912). When the attributes of the table are independent and only in this case it is zero. It is $+1$ only if b and c are both zero (i.e. if the two rankings agree in all pairs and hence are identical) and -1 only if a and d are both zero (i.e. if the two rankings disagree in all pairs and one is the reverse of the other). In this latter property of attaining unity only when diagonally opposite cells are both empty it differs from two other coefficients proposed by Yule.*

Thus partial r as defined is a measure of the association of agreements of the rankings 2 and 3 when compared in pairs with ranking 1. From this viewpoint it will be seen that the use of the word 'partial' conforms to that in the theories of association and correlation. The partial associations of A and B in a third population containing C and γ are those of AC and BC or of $A\gamma$ and $B\gamma$. The partial association of ranks is that of pairs of agreements of 21 and 31. The partial correlation of 2 and 3 when 1 is constant is paralleled by the partial rank correlation of 2 and 3 when 1 is in the natural order 1, ..., n. If partial r is unity in absolute value the rankings coincide or one is the reverse of the other, whether they agree with ranking 1 or not, so that the coefficient fulfils its proper function of measuring the relationship between 2 and 3 independently of the influence of 1.

* Yule himself arrived at the coefficient by considering product-moments in a 2×2 table. Karl Pearson & Heron (1913) mistook Yule's intention and thought the coefficient was proposed as an estimate of the correlation in a normal population whose frequencies were given by a double dichotomy in the 2×2 table. Their long memoir is mainly devoted to advocating the alternative merits of tetrachoric r. The two things, as Yule has emphasized, are quite different. I mention the point to make it clear that the Pearson-Heron criticisms of Yule's coefficient, even if not misfounded, do not affect the above work, since I use the coefficient purely as a measure of association in the fourfold table.

8. We may establish the remarkable result

$$\tau_{23.1} = \frac{\tau_{23} - \tau_{12}\tau_{13}}{(1-\tau_{12}^2)^{\frac{1}{2}}(1-\tau_{13}^2)^{\frac{1}{2}}}. \tag{15}$$

In fact, from expression (12) we see that

$$\tau_{12} = \frac{(a+b)-(c+d)}{N},$$

$$\tau_{13} = \frac{(a+c)-(b+d)}{N},$$

$$\tau_{23} = \frac{(a+d)-(b+c)}{N}.$$

Remembering that $N = a+b+c+d$ we have

$$1-\tau_{12}^2 = \frac{4}{N^2}(a+b)(c+d),$$

$$1-\tau_{13}^2 = \frac{4}{N^2}(a+c)(b+d),$$

$$\begin{aligned}\tau_{23}-\tau_{12}\tau_{13} &= \frac{1}{N^2}[(a+b+c+d)\{(a+d)-(b+c)\} \\ &\qquad -\{(a+b)-(c+d)\}\{(a+c)-(b+d)\}] \\ &= \frac{4}{N^2}(ad-bc).\end{aligned}$$

Equation (15) follows at once from the definition of partial τ in equation (13).

The appearance of the product-moment type of relation between total and partial correlations is surprising. There was no reason to expect that partial τ, which is a pure function of disarrangements in rankings and is not expressible algebraically in terms of the ranks, should bear any analogy with the partial correlation of variates; but since it does so, we are evidently fortified in regarding partial τ as a convenient measure of rank correlation.

Example. Ten men are ranked according to (1) intelligence, (2) mathematical ability, (3) musical ability. The rankings are:

1:	1	2	3	4	5	6	7	8	9	10
2:	1	4	5	6	2	7	3	9	8	10
3:	4	1	3	5	2	6	7	10	9	8

It will be found that $\tau_{12} = 0{\cdot}644$, $\tau_{13} = 0{\cdot}644$, $\tau_{23} = 0{\cdot}555$. Thus mathematical and musical ability are positively correlated. The question is, can this correlation be attributed to the correlation of both with intelligence?

We find

$$\tau_{23.1} = \frac{0{\cdot}555-(0{\cdot}644)^2}{1-(0{\cdot}644)^2}$$

$$= 0{\cdot}24.$$

The conclusion would be that although part of the total correlation is due to the correlation of both with intelligence, part of it is not. We cannot attribute the whole of the observed (total) correlation between mathematical and musical ability to the interference of common correlation with intelligence.

9. The same methods are immediately capable of extension to paired comparisons. In fact the array of type (10) is the array of paired comparisons for all the possible pairs of ranks and when there are no constraints of the ranking character (i.e. such that if $A \to B$ and $B \to C$ then must $A \to C$) the coefficient τ becomes a measure of agreement in paired comparisons (cf. Kendall & Babington Smith, 1940). We could then construct measures of partial agreement by the same formulae in cases where it was suspected that there were mutual influences at work between three observers, as for instance if it were suspected that community of preference between two children was due to community of both with one of the parents.

10. In conclusion, it may perhaps be desirable to point out that although partial τ is defined by equation (14) in terms of χ^2 for a fourfold table, its significance cannot on that account be tested in the Type III χ^2 distribution with one degree of freedom. I have not yet succeeded in finding expressions for the sampling distribution of partial τ but it seems clear that the Type III distribution will not be reproduced in the ranking case, at least without some substantial modification, even when the rankings are independent; for there exist correlations between the signs given by any ranking in the recorded order. If, for example, (12) and (23) are positive so must be (13), whereas if (13) and (23) are positive (12) may be either positive or negative. The units in the fourfold table cannot therefore be regarded as allocated at random and the type III distribution will probably not hold. I hope to return to this subject on a later occasion.

REFERENCES

KENDALL, M. G. (1938). A new measure of rank correlation. *Biometrika*, **30**, 81.

KENDALL, M. G. & BABINGTON SMITH, B. (1940). On the method of paired comparisons. *Biometrika*, **31**, 324.

PEARSON, KARL & HERON, D. (1913). On theories of association. *Biometrika*, **9**, 159.

YULE, G. UDNY (1912). On the methods of measuring the association between two attributes. *J.R. Statist. Soc.* **75**, 579.

INEQUALITIES FOR MULTIVARIATE FREQUENCY DISTRIBUTIONS

By C. E. V. LESER

Given a frequency distribution y of a single variate x, with arithmetic mean $\bar{x} = 0$ and with standard deviation σ, Tchebycheff's (1867) well-known inequality presents a lower limit for the ratio of the frequency of all values of x between $-\lambda\sigma$ and $\lambda\sigma$ to the total frequency, where $\lambda \geqq 1$. In the case of the special class of frequency distributions for which $y(x)+y(-x)$ is a non-increasing function of $|x|$, this inequality can be substantially improved to another limit which applies to all positive values of λ, as already Gauss (1823) has proved. Various authors* have generalized these theorems by modifying the assumptions made with regard to the frequency distribution, by introducing moments of higher order than the second or by extending some of the results to bivariate functions. In the present analysis only moments of second order are used, but the results apply to frequency functions of any number of variates.

Suppose $y(x_1, \ldots, x_n)$ to be a frequency distribution of n variates, with total frequency equal to unity, arithmetic mean at the origin and with standard deviations $\sigma_1, \ldots, \sigma_n$. Let P be the frequency of all combinations of $x_1, \ldots, x_n$ for which

$$\left(\frac{x_1}{\lambda_1\sigma_1}\right)^2 + \ldots + \left(\frac{x_n}{\lambda_n\sigma_n}\right)^2 \leqq n.$$

Write $$\lambda_0 = \sqrt{\frac{n}{1/\lambda_1^2 + \ldots + 1/\lambda_n^2}}, \quad \sigma_0 = \sqrt{\frac{n}{1/\sigma_1^2 + \ldots + 1/\sigma_n^2}},$$

so that λ_0^2 and σ_0^2 are the harmonic averages of $\lambda_1^2, \ldots, \lambda_n^2$ and $\sigma_1^2, \ldots, \sigma_n^2$ respectively. We also write

$$\mu_i = \frac{\lambda_i}{\lambda_0}, \quad \tau_i = \frac{\sigma_i}{\sigma_0} \quad (i = 1, \ldots, n),$$

$$R = \sqrt{\left\{\left(\frac{x_1}{\mu_1\tau_1}\right)^2 + \ldots + \left(\frac{x_n}{\mu_n\tau_n}\right)^2\right\}} = \lambda_0\sigma_0\sqrt{\left\{\left(\frac{x_1}{\lambda_1\sigma_1}\right)^2 + \ldots + \left(\frac{x_n}{\lambda_n\sigma_n}\right)^2\right\}},$$

so that P is the frequency of all values of $x_1, \ldots, x_n$ for which $R/\sqrt{n} \leqq \lambda_0\sigma_0$. Furthermore, we define $A(R_0)$ as the average value of y for all those values of $x_1, \ldots, x_n$ for which R has a fixed value R_0, and therefore

$$A(R_0) = \frac{\int_{R=R_0} \ldots \int y\, dx_1 \ldots dx_n}{\int_{R=R_0} \ldots \int dx_1 \ldots dx_n}.$$

* E.g. K. Pearson (1919), B. H. Camp (1922), S. Narumi (1923), C. D. Smith (1930), P. O. Berge (1937).

We can now state the following:

THEOREM. Assuming the frequency distribution to be such that $A(R)$ is a non-increasing function of R for $R/\sqrt{n} \leqq \kappa\sigma_0$, we have one of the following three sets of inequalities, according to the value of κ:

I. $\kappa \leqq 1$:

(i) $\lambda_0 \leqq 1$ $\qquad P \geqq 0,$

(ii) $1 \leqq \lambda_0$ $\qquad P \geqq 1 - \frac{1}{\lambda_0^2}.$

II. $1 \leqq \kappa \leqq \left(\frac{n+2}{n}\right)^{\frac{1}{2}}$:

(i) $\lambda_0 \leqq \left(\frac{2}{n+2}\right)^{1/n} \kappa$ $\qquad P \geqq \frac{n+2}{2}\left(1 - \frac{1}{\kappa^2}\right)\left(\frac{\lambda_0}{\kappa}\right)^n,$

(ii) $\left(\frac{2}{n+2}\right)^{1/n} \kappa \leqq \lambda_0 \leqq \kappa$ $\qquad P \geqq 1 - \frac{1}{\kappa^2},$

(iii) $\kappa \leqq \lambda_0$ $\qquad P \geqq 1 - \frac{1}{\lambda_0^2}.$

III. $\left(\frac{n+2}{n}\right)^{\frac{1}{2}} \leqq \kappa$:

(i) $\lambda_0 \leqq \left(\frac{2}{n+2}\right)^{1/n}\left(\frac{n+2}{n}\right)^{\frac{1}{2}}$ $\qquad P \geqq \left(\frac{n}{n+2}\right)^{\frac{1}{2}n} \lambda_0^n,$

(ii) $\left(\frac{2}{n+2}\right)^{1/n}\left(\frac{n+2}{n}\right)^{\frac{1}{2}} \leqq \lambda_0 \leqq \left(\frac{2}{n+2}\right)^{1/n} \kappa$ $\qquad P \geqq 1 - \left(\frac{2}{n+2}\right)^{2/n} \frac{1}{\lambda_0^2},$

(iii) $\left(\frac{2}{n+2}\right)^{1/n} \kappa \leqq \lambda_0 \leqq \kappa$ $\qquad P \geqq 1 - \frac{1}{\kappa^2},$

(iv) $\kappa \leqq \lambda_0$ $\qquad P \geqq 1 - \frac{1}{\lambda_0^2}.$

Proof. We introduce the new system of co-ordinates

$$\begin{aligned}
x_1 &= \mu_1 \tau_1 R \cos t_1 \cos t_2 \cos t_3 \dots \cos t_{n-1},\\
x_2 &= \mu_2 \tau_2 R \sin t_1 \cos t_2 \cos t_3 \dots \cos t_{n-1},\\
x_3 &= \mu_3 \tau_3 R \quad \sin t_2 \cos t_3 \dots \cos t_{n-1},\\
&\dots\dots\dots

We also write $$s = \sigma_0 \sqrt{n},$$

and we are going to use the following abbreviating symbol

$$\int_T f(R,t)\,dt = \int_{t_1=0}^{2\pi} \int_{t_2=-\frac{1}{2}\pi}^{\frac{1}{2}\pi} \dots \int_{t_{n-1}=-\frac{1}{2}\pi}^{\frac{1}{2}\pi} f(R,t_1,\dots,t_{n-1})\,dt_1 \dots dt_{n-1}.$$

Then $$\int_{R=0}^{\infty} \int_T R^{n-1} z(R,t)\,dR\,dt = 1,$$

$$\int_{R=0}^{\lambda_0 s} \int_T R^{n-1} z(R,t)\,dR\,dt = P,$$

and as $A(R) = \text{const.} \int_T z(R,t)\,dt$, we have the condition that $\int_T z(R,t)\,dt$ is a non-increasing function of R for $R \leqq \kappa s$. Furthermore, let us write

$$G = \int_T z(\lambda_0 s, t)\,dt,$$

and later on $$u = \frac{G}{n} s^n.$$

Now start from the equation

$$\left(\frac{\sigma_1}{\mu_1 \tau_1}\right)^2 + \dots + \left(\frac{\sigma_n}{\mu_n \tau_n}\right)^2$$

$$= \int_{x_1=-\infty}^{\infty} \dots \int_{x_n=-\infty}^{\infty} \left[\left(\frac{x_1}{\mu_1 \tau_1}\right)^2 + \dots + \left(\frac{x_n}{\mu_n \tau_n}\right)^2\right] y(x_1,\dots,x_n)\,dx_1,\dots,dx_n$$

which can be written

$$(\lambda_0 \sigma_0)^2 \left(\frac{1}{\lambda_1^2} + \dots + \frac{1}{\lambda_n^2}\right) = \int_{R=0}^{\infty} \int_T R^{n+1} z(R,t)\,dR\,dt$$

or $$s^2 = I_1 + I_2,$$

$$I_1 = \int_{R=0}^{\lambda_0 s} \int_T R^{n+1} z(R,t)\,dR\,dt, \quad I_2 = \int_{R=\lambda_0 s}^{\infty} \int_T R^{n+1} z(R,t)\,dR\,dt.$$

According to the value of κs, three cases have to be distinguished.

(*a*) $\kappa s \geqq \left[(\lambda_0 s)^n + \frac{n}{G}(1-P)\right]^{1/n}$ or $P \geqq 1 - (\kappa^n - \lambda_0^n)\,u$.

For $R \leqq \lambda_0 s$ $$\int_T z(R,t)\,dt \geqq G.$$

Hence $$I_1 \geqq G \int_{R=0}^{\lambda_0 s} R^{n+1}\,dR,$$

and incidentally $$P \geqq G \int_{R=0}^{\lambda_0 s} R^{n-1}\,dR = \lambda_0^n u.$$

To obtain a limit for I_2, we make use of the equation

$$\int_{R=\lambda_0 s}^{\infty}\int_T R^{n-1}z(R,t)\,dR\,dt = 1-P = G\int_{R=\lambda_0 s}^{\left[(\lambda_0 s)^n+\frac{n}{G}(1-P)\right]^{1/n}} R^{n-1}dR$$

or

$$\int_{R=\left[(\lambda_0 s)^n+\frac{n}{G}(1-P)\right]^{1/n}}^{\infty}\int_T R^{n-1}z(R,t)\,dR\,dt = \int_{R=\lambda_0 s}^{\left[(\lambda_0 s)^n+\frac{n}{G}(1-P)\right]^{1/n}} R^{n-1}\left[G-\int_T z(R,t)\,dt\right]dR.$$

As the integrands are nowhere negative and R nowhere smaller in the integral on the left than in the integral on the right, we can write

$$\int_{R=\left[(\lambda_0 s)^n+\frac{n}{G}(1-P)\right]^{1/n}}^{\infty}\int_T R^{n+1}z(R,t)\,dR\,dt \geqq \int_{R=\lambda_0 s}^{\left[(\lambda_0 s)^n+\frac{n}{G}(1-P)\right]^{1/n}} R^{n+1}\left[G-\int_T z(R,t)\,dt\right]dR,$$

$$I_2 \geqq G\int_{R=\lambda_0 s}^{\left[(\lambda_0 s)^n+\frac{n}{G}(1-P)\right]^{1/n}} R^{n+1}\,dR,$$

and therefore

$$s^2 \geqq G\int_{R=0}^{\left[(\lambda_0 s)^n+\frac{n}{G}(1-P)\right]^{1/n}} R^{n+1}\,dR = \frac{[G(\lambda_0 s)^n+n(1-P)]^{(n+2)/n}}{(n+2)\,G^{2/n}},$$

$$n^{2/n}(n+2)\,u^2 \geqq [n(\lambda_0^n u+1-P)]^{(n+2)/n},$$

$$P \geqq 1+\lambda_0^n u-\left(\frac{n+2}{n}\right)^{n/(n+2)}u^{2/(n+2)}.$$

(*b*) $\lambda_0 s \leqq \kappa s \leqq \left[(\lambda_0 s)^n+\frac{n}{G}(1-P)\right]^{1/n}$ or $P \leqq 1-(\kappa^n-\lambda_0^n)\,u, \quad \lambda_0 \leqq \kappa.$

As before, $I_1 \geqq G\int_{R=0}^{\lambda_0 s} R^{n+1}\,dR$ and $P \geqq \lambda_0^n u.$

Furthermore, we have the equation

$$\int_{R=\lambda_0 s}^{\infty}\int_T R^{n-1}z(R,t)\,dR\,dt = 1-P = G\int_{R=\lambda_0 s}^{\kappa s} R^{n-1}dR+\left[1-P-\frac{G}{n}(\kappa^n-\lambda_0^n)\,s^n\right],$$

$$\int_{R=\kappa t}^{\infty}\int_T R^{n-1}z(R,t)\,dR\,dt = \int_{R=\lambda_0 s}^{\kappa s} R^{n-1}\left[G-\int_T z(R,t)\,dt\right]dR + \left[1-P-\frac{G}{n}(\kappa^n-\lambda_0^n)\,s^n\right],$$

and, as before, it follows that

$$\int_{R=\kappa s}^{\infty}\int_T R^{n+1}z(R,t)\,dR\,dt \geqq \int_{R=\lambda_0 s}^{\kappa s} R^{n+1}\left[G-\int_T z(R,t)\,dt\right]dR + (\kappa s)^2\left[1-P-\frac{G}{n}(\kappa^n-\lambda_0^n)\,s^n\right],$$

$$I_2 \geqq G\int_{R=\lambda_0 s}^{\kappa s} R^{n+1}dR + (\kappa s)^2\left[1-P-\frac{G}{n}(\kappa^n-\lambda_0^n)\,s^n\right].$$

Hence

$$s^2 \geqq G\int_{R=0}^{\kappa s} R^{n+1}dR + (\kappa s)^2\left[1-P-\frac{G}{n}(\kappa^n-\lambda_0^n)\,s^n\right],$$

$$1 \geqq \frac{n}{n+2}\kappa^{n+2}u + \kappa^2[1-P-(\kappa^n-\lambda_0^n)\,u],$$

$$P \geqq \left(1-\frac{1}{\kappa^2}\right)+\left(\lambda_0^n-\frac{2}{n+2}\kappa^n\right)u.$$

(*c*) $\kappa s \leqq \lambda_0 s$ or $\lambda_0 \geqq \kappa$.

In this case, I_1 need not be larger than 0, but

$$I_2 \geqq (\lambda_0 s)^2(1-P),$$

$$s^2 \geqq (\lambda_0 s)^2(1-P).$$

Therefore, if $\lambda_0 \geqq 1$, $P \geqq 1-\frac{1}{\lambda_0^2}$, and if $\lambda_0 \leqq 1$, no significant limit exists.

For $\lambda_0 \geqq \kappa$, the problem has thus been solved, but the case $\lambda_0 \leqq \kappa$ is more complicated; summarizing the results of (*a*) and (*b*), we have two alternative sets of inequalities. Write

$$f_1(u) = \lambda_0^n u, \quad f_2(u) = 1-(\kappa^n-\lambda_0^n)\,u,$$

$$f_3(u) = 1+\lambda_0^n u-\left(\frac{n+2}{n}\right)^{n/(n+2)}u^{2/(n+2)},$$

$$f_4(u) = \left(1-\frac{1}{\kappa^2}\right)+\left(\lambda_0^n-\frac{2}{n+2}\kappa^n\right)u.$$

Then either $P \geqq f_1, \quad P \geqq f_2, \quad P \geqq f_3,$

or $P \geqq f_1, \quad P \leqq f_2, \quad P \geqq f_4.$

For different frequency distributions, G and therefore u may assume any non-negative value which is compatible with the condition $P \leqq 1$. We have to find the effective lower limit for P as a function of u and its minimum value which, as easily seen, is only larger than 0, if $f_4(0) > 0$ or, what is the same thing, $\kappa > 1$.

f_1, f_2, f_4 are straight lines. As

$$\frac{df_3}{du} = \lambda_0^n - \frac{2}{n+2}\left(\frac{n+2}{n}\right)^{n/(n+2)}u^{-n/(n+2)}$$

f_3 has a minimum when

$$u_3^{(\min.)} = \frac{n+2}{n}\left(\frac{2}{n+2}\right)^{(n+2)/n}\frac{1}{\lambda_0^{n+2}} = \frac{2}{n}\left(\frac{2}{n+2}\right)^{2/n}\frac{1}{\lambda_0^{n+2}}$$

and $$f_3^{(\min.)} = 1+\frac{2}{n}\left(\frac{2}{n+2}\right)^{2/n}\frac{1}{\lambda_0^2}-\frac{n+2}{n}\left(\frac{2}{n+2}\right)^{2/n}\frac{1}{\lambda_0^2} = 1-\left(\frac{2}{n+2}\right)^{2/n}\frac{1}{\lambda_0^2}.$$

Indicating by u_{ij}, f_{ij} the co-ordinates of the point in which the curves f_i and f_j intersect, we find that

$$u_{12} = \frac{1}{\kappa^n}, \qquad f_{12} = \left(\frac{\lambda_0}{\kappa}\right)^n,$$

$$u_{13} = \left(\frac{n}{n+2}\right)^{\frac{1}{2}n}, \qquad f_{13} = \left(\frac{n}{n+2}\right)^{\frac{1}{2}n}\lambda_0^n,$$

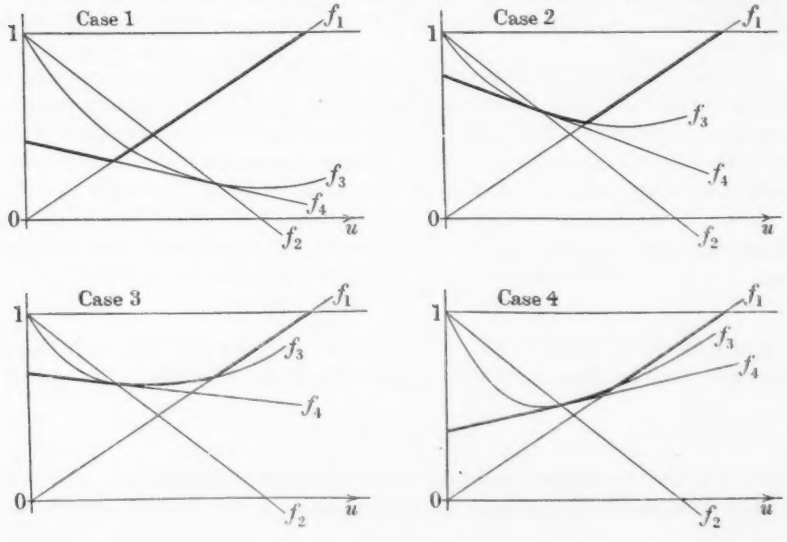
$$u_{14} = \frac{n+2}{2}\left(1-\frac{1}{\kappa^2}\right)\frac{1}{\kappa^n}, \quad f_{14} = \frac{n+2}{2}\left(1-\frac{1}{\kappa^2}\right)\left(\frac{\lambda_0}{\kappa}\right)^n.$$

f_2, f_3 and f_4 have a point (u_{234}, f_{234}) in common in which f_4 is the tangent to f_3:

$$u_{234} = \frac{n+2}{n}\frac{1}{\kappa^{n+2}}, \quad f_{234} = 1-\frac{n+2}{n}\frac{\kappa^n-\lambda_0^n}{\kappa^{n+2}}.$$

It is also seen that the sign of both expressions $\dfrac{df_4}{du}$ and $u_{234}-u_3^{(\min.)}$ is positive or negative according to whether $\lambda_0 \gtrless \left(\dfrac{2}{n+2}\right)^{1/n}\kappa$.

We have therefore four possibilities. In the following diagrams, the heavy



line shows the effective lower limit for P as a function of u (the upper limit is always equal to 1, of course), and its minimum gives the general lower limit for P.

In these four cases, the following results are obtained:

(1)		$u_{234} \geqq u_{12}$,		$P \geqq f_{14}$,
(2)	$\frac{df_4}{du} \leqq 0$,	$u_{234} \leqq u_{12}$,	$u_3^{(\min.)} \geqq u_{13}$	$P \geqq f_{13}$,
(3)		$u_{234} \leqq u_{12}$,	$u_3^{(\min.)} \leqq u_{13}$	$P \geqq f_3^{(\min.)}$
(4)	$\frac{df_4}{du} \geqq 0$			$P \geqq f_4(0)$,

and by substituting the proper values:

(1)	$1 \leqq \kappa \leqq \left(\frac{n+2}{n}\right)^{\frac{1}{2}}$,	$\lambda_0 \leqq \left(\frac{2}{n+2}\right)^{1/n} \kappa$	$P \geqq \frac{n+2}{2}\left(1-\frac{1}{\kappa^2}\right)\left(\frac{\lambda_0}{\kappa}\right)^n$,
(2)	$\kappa \geqq \left(\frac{n+2}{n}\right)^{\frac{1}{2}}$,	$\lambda_0 \leqq \left(\frac{2}{n+2}\right)^{1/n}\left(\frac{n+2}{n}\right)^{\frac{1}{2}}$	$P \geqq \left(\frac{n}{n+2}\right)^{\frac{1}{2}n} \lambda_0^n$
(3)		$\left(\frac{2}{n+2}\right)^{1/n} \kappa \geqq \lambda_0 \geqq \left(\frac{2}{n+2}\right)^{1/n}\left(\frac{n+2}{n}\right)^{\frac{1}{2}}$	$P \geqq 1-\left(\frac{2}{n+2}\right)^{2/n}\frac{1}{\lambda_0^2}$,
(4)	$\kappa \geqq 1$,	$\left(\frac{2}{n+2}\right)^{1/n} \kappa \leqq \lambda_0 \leqq \kappa$	$P \geqq 1-\frac{1}{\kappa^2}$.

In addition, we know that for $\lambda_0 \geqq \kappa$, $\lambda_0 \geqq 1$: $P \geqq 1-\frac{1}{\lambda_0^2}$. By rearranging these inequalities, we can bring them into the form in which they were given in the theorem, which is therefore proved.

It may be remarked that κ depends on the ratios between any two of the quantities $\lambda_1, \ldots, \lambda_n$, but it remains the same if $\lambda_1, \ldots, \lambda_n$ change in the same proportion.

Let us consider the sets of inequalities I, II and III of the theorem separately. The most important case in which set I is relevant occurs when nothing about the frequency distribution is known except the averages and standard deviations of the variables, in which case we have to put $\kappa = 0$. It provides a generalization of Tchebycheff's inequality which is obtained in the special case $n = 1$ in which $\lambda_0 = \lambda$.

If $n = 2$, the inequality refers to the frequency of all points lying inside the ellipse which has the axes $\lambda_1 \sqrt{2}\sigma_1$, $\lambda_2 \sqrt{2}\sigma_2$, and it can be written in the following way:

$$\frac{1}{\lambda_1^2}+\frac{1}{\lambda_2^2} \leqq 2 : P \geqq 1-\frac{1}{2}\left(\frac{1}{\lambda_1^2}+\frac{1}{\lambda_2^2}\right).$$

It is interesting to compare this limit with the one given by Berge (1937) for a rectangle which has its corners in the points with the co-ordinates $(\pm\lambda\sigma_1, \pm\lambda\sigma_2)$:

$$P \geqq 1-\frac{1+\sqrt{(1-r^2)}}{\lambda^2},$$

where r is the correlation coefficient of the frequency distribution. If $r = 0$, Berge's limit equals $1-2/\lambda^2$ and is therefore equal to the limit we get for the

ellipse with the axes $\lambda\sigma_1$, $\lambda\sigma_2$ which is inscribed to the rectangle. On the other hand, if $r = \pm 1$, Berge's limit becomes equal to $1 - 1/\lambda^2$, and we have to choose the circumscribed ellipse with the axes $\lambda\sqrt{2}\sigma_1$, $\lambda\sqrt{2}\sigma_2$ to obtain the same limit. Hence, the limit given here is certainly not inferior to the limit given by Berge, if the two variates are independent, but it is not superior to this limit, if there is a perfect correlation.

Set II is perhaps of more theoretical than practical interest. An intermediate value of κ may, however, be realized, if there is sufficient information about the frequencies inside, but not outside, a certain n-dimensional interval. The inequalities correspond to those given by Narumi (1923) for frequency distributions of one variate only, but having a more general meaning in so far as P may refer to multiples of other quantities besides the standard deviation. The first inequality seems the most interesting one of this set; it reduces to a special case of Narumi's inequality if we put $n = 1$:

$$\lambda \leqq \tfrac{2}{3}\kappa : P \geqq \frac{3}{2}\left(1 - \frac{1}{\kappa^2}\right)\frac{\lambda}{\kappa},$$

and for $n = 2$ it may be written

$$\lambda_0 \leqq \frac{\kappa}{\sqrt{2}} : P \geqq 2\left(1 - \frac{1}{\kappa^2}\right)\left(\frac{\lambda_0}{\kappa}\right)^2,$$

or

$$\frac{1}{\lambda_1^2} + \frac{1}{\lambda_2^2} \geqq \frac{4}{\kappa^2} : P \geqq 4\left(1 - \frac{1}{\kappa^2}\right)\frac{1}{\kappa^2(1/\lambda_1^2 + 1/\lambda_2^2)}.$$

Set III generalizes Gauss's inequalities. It reduces to the first two inequalities for those values of $\lambda_1, \ldots, \lambda_n$ (if any) for which $\kappa = \infty$. This is the case for all values of $\lambda_1, \ldots, \lambda_n$, if $y(hx_1, \ldots, hx_n) + y(-hx_1, \ldots, -hx_n)$ is a non-increasing function of $|h|$ for any fixed values of $x_1, \ldots, x_n$; especially if the function y decreases monotonically along every straight line radiating from the origin.

The assumption $\kappa = \infty$ will be made throughout the following analysis. Gauss's formulae are obtained by substituting $n = 1$:

$$\lambda \leqq \frac{2}{\sqrt{3}} : P \geqq \frac{\lambda}{\sqrt{3}},$$

$$\lambda \geqq \frac{2}{\sqrt{3}} : P \geqq 1 - \frac{4}{9\lambda^2}.$$

In the case $n = 2$, the inequalities are also greatly simplified:

$$\lambda_0 \leqq 1 : P \geqq \frac{\lambda_0^2}{2},$$

$$\lambda_0 \geqq 1 : P \geqq 1 - \frac{1}{2\lambda_0^2},$$

or

$$\frac{1}{\lambda_1^2} + \frac{1}{\lambda_2^2} \geqq 2 : P \geqq \frac{1}{1/\lambda_1^2 + 1/\lambda_2^2},$$

$$\frac{1}{\lambda_1^2} + \frac{1}{\lambda_2^2} \leqq 2 : P \geqq 1 - \frac{1}{4}\left(\frac{1}{\lambda_1^2} + \frac{1}{\lambda_2^2}\right).$$

Returning to the case of an unspecified value of n, we shall often be interested in a limit for the frequency of all values of the standardized or original variables for which the distance from the origin is no more than a certain value, i.e. either

$$\frac{x_1^2}{\sigma_1^2}+\ldots+\frac{x_n^2}{\sigma_n^2}\leqq\mu^2,$$

or

$$x_1^2+\ldots+x_n^2\leqq\rho^2.$$

In the first problem, we have

$$\lambda_1=\ldots=\lambda_n=\lambda_0=\frac{\mu}{\sqrt{n}},$$

and our inequalities can therefore be written

$$\mu^2\leqq 2^{2/n}(n+2)^{(n-2)/n}: P\geqq\frac{\mu^n}{(n+2)^{\frac{1}{2}n}},$$

$$\mu^2\geqq 2^{2/n}(n+2)^{(n-2)/n}: P\geqq 1-\left(\frac{2}{n+2}\right)^{2/n}\frac{n}{\mu^2}.$$

In the second problem

$$\lambda_1\sigma_1=\ldots=\lambda_n\sigma_n=\frac{\rho}{\sqrt{n}},\quad \lambda_0=\frac{\rho}{\sqrt{(\sigma_1^2+\ldots+\sigma_n^2)}}.$$

Hence, if

$$\rho^2\leqq 2^{2/n}(n+2)^{(n-2)/n}\frac{\sigma_1^2+\ldots+\sigma_n^2}{n}: P\geqq\left[\frac{n\rho^2}{(n+2)(\sigma_1^2+\ldots+\sigma_n^2)}\right]^{\frac{1}{2}n},$$

$$\rho^2\geqq 2^{2/n}(n+2)^{(n-2)/n}\frac{\sigma_1^2+\ldots+\sigma_n^2}{n}: P\geqq 1-\left(\frac{2}{n+2}\right)^{2/n}\frac{\sigma_1^2+\ldots+\sigma_n^2}{\rho^2}.$$

Again, the insertion of special values for n will simplify the formulae.

Finally, it may be at least of theoretical interest to consider the generalized Gauss limit, as obtained for $\kappa=\infty$, as a function of n for fixed values of λ_0, and to compare it with the generalized Tchebycheff limit which is obtained for $\kappa\leqq 1$ and is independent of n. For some particular values of λ_0 and n, this is done in the following table:

n \ λ_0	1	2	3
		Gauss limit	
1	0·577	0·889	0·951
2	0·500	0·875	0·944
3	0·457	0·864	0·940
4	0·423	0·856	0·936
		Tchebycheff limit	
Any value	0	0·750	0·889

Furthermore, since for $\lambda_0 \leqq \left(\frac{2}{n+2}\right)^{1/n}\left(\frac{n+2}{n}\right)^{\frac{1}{2}}$,

$$\lim_{n\to\infty}\left(\frac{n}{n+2}\right)^{\frac{1}{2}n}\lambda_0^n \leqq \lim_{n\to\infty}\frac{2}{n+2} = 0,$$

and since

$$\lim_{n\to\infty}\left[1-\left(\frac{2}{n+2}\right)^{2/n}\frac{1}{\lambda_0^2}\right] = 1-\frac{1}{\lambda_0^2},$$

it is seen that with an increasing number of variates, the Gauss limit loses gradually its superiority over the Tchebycheff limit and the difference between the two limits tends to vanish.

REFERENCES

BERGE, P. O. (1937). A note on a form of Tchebycheff's theorem for two variables. *Biometrika*, **29**, 405–6.

CAMP, B. H. (1922). A new generalization of Tchebycheff's statistical inequality. *Bull. Amer. Math. Soc.* **28**, 427–32.

GAUSS (1823). *Theoria combinationis observationum erroribus minimis obnoxiae.* Göttingen, pp. 9–12.

NARUMI, S. (1923). On further inequalities with possible applications to problems in the theory of probability. *Biometrika*, **15**, 245–53.

PEARSON, K. (1919). On generalized Tchebycheff theorems in the mathematical theory of statistics. *Biometrika*, **12**, 284–96.

SMITH, C. D. (1930). On generalized Tchebycheff inequalities in mathematical statistics. *Amer. J. Math.* **52**, 109–26.

TCHEBYCHEFF (1867). Des valeurs moyennes. *J. Math. pures appl.* ser. 2, **12**, 177–84.

THE MODE AND MEDIAN OF A NEARLY NORMAL DISTRIBUTION WITH GIVEN CUMULANTS

By J. B. S. HALDANE, F.R.S.

In the early years of biometry the mode and median of a distribution, and especially the latter, were regarded as being almost as important as the mean. Later Pearson and others developed the method of moments, and recently the cumulants, which are readily derived from the moments, have been widely used. Pearson (1895 and after) discussed the relation of the mean, mode and median in some of the skew frequency curves which he invented. He found empirically that for skew curves of Type III, namely,

$$y = y_0\left(1+\frac{x}{a}\right)^p e^{-px/a},$$

when p is positive, Mode − Mean = 3 (Median − Mean) approximately. By fitting for a series of integral values of p he found

$$\frac{\text{Median}-\text{Mode}}{\text{Mean}-\text{Mode}} = 0{\cdot}6691+0{\cdot}0094p^{-1}.$$

Some later writers have taken the trisection as a general law. But as we shall see, it does not always hold, even approximately. So far as I know, general expressions for the mode and median in terms of the cumulants have not been given, nor have the conditions been stated under which Pearson's rule holds approximately.

Consider a variate X, with distribution $df = F(X)\,dX$, and cumulants $\kappa_1 = m$, $\kappa_2 = \sigma^2$, $\kappa_3, \ldots, \kappa_r, \ldots$.

The algebra is simplified if we make the transformation $x = (X-m)/\sigma$, so that x has mean zero, and unit standard deviation, its cumulants being $0, 1, \gamma_1, \gamma_2, \ldots, \gamma_r, \ldots$, where $\gamma_{r-2} = \kappa_r \kappa_2^{-\frac{1}{2}r}$. Thus γ_r is the rth measure of the deviation from normality of the distribution of X. Now γ_r may become infinite for all values of r, or for all even values of r, above a certain value. It may diminish indefinitely or remain small. In other cases it falls with r at first, but then increases without limit. Thus for Pearson's Type III distribution, whose equation has been given above,

$$\kappa_r = (r-1)!\,(p+1)\left(\frac{a}{p}\right)^r, \quad \text{and} \quad \gamma_r = (r+1)!\,(p+1)^{-\frac{1}{2}r}.$$

And in the case of any estimate of a statistic, such as the mean, variance, standard deviation, or skewness, based on a sample of n members, $\gamma_r = O(n^{-\frac{1}{2}r})$. We shall not discuss the convergence of the series obtained later. But it is worth

noting that expansions in Hermitian polynomials are often satisfactory asymptotic expansions, even if they diverge.

Let $df = f(x)\,dx$ be the distribution of x. Then it is known that we may write, symbolically,*

$$f(x) = \exp\left[-\frac{\gamma_1}{3!}\left(\frac{d}{dx}\right)^3 + \frac{\gamma_2}{4!}\left(\frac{d}{dx}\right)^4 - \ldots + \frac{\gamma_{r-2}}{r!}\left(\frac{-d}{dx}\right)^r \ldots\right]\frac{e^{-\frac{1}{2}x^2}}{\sqrt{(2\pi)}}.$$

Expanding in Hermitian polynomials $H_r(x) = e^{\frac{1}{2}x^2}\frac{d^r}{dx^r}e^{-\frac{1}{2}x^2}$, we have

$$f(x) = \frac{e^{-\frac{1}{2}x^2}}{\sqrt{(2\pi)}}\left[1 - \frac{\gamma_1}{3!}H_3(x) + \frac{\gamma_2}{4!}H_4(x) - \frac{\gamma_3}{5!}H_5(x) + \frac{10\gamma_1^2+\gamma_4}{6!}H_6(x) - \frac{35\gamma_1\gamma_2+\gamma_5}{7!}H_7(x) + \ldots\right].$$

In the special case where $\gamma_r = O(n^{-\frac{1}{2}r})$ we have

$$f(x) = \frac{e^{-\frac{1}{2}x^2}}{\sqrt{(2\pi)}}\left[1 - \frac{\gamma_1}{6}H_3(x) + \frac{\gamma_2}{24}H_4(x) - \frac{\gamma_3}{120}H_5(x) + \frac{\gamma_1^2}{72}H_6(x) - \frac{\gamma_1\gamma_2}{144}H_7(x) - \frac{\gamma_1^3}{1296}H_9(x) + O(n^{-2})\right].$$

To find the mode, we put $\frac{d}{dx}f(x) = 0$. Hence

$$H_1(x) - \frac{\gamma_1}{6}H_4(x) + \frac{\gamma_2}{24}H_5(x) + \ldots = 0.$$

The terms needed to find the root of this equation with the smallest absolute value depend, of course, on the behaviour of the cumulants as r increases. If $|\gamma_2|$, $|\gamma_3|$, etc., are not substantially less than $|\gamma_1|$, but all are small, so that powers and products can be neglected, then:

$$x = -\frac{\gamma_1}{2} + \frac{\gamma_3}{8} - \frac{\gamma_5}{48} + \frac{\gamma_7}{576} - \ldots.$$

Where $\gamma_r = O(n^{-\frac{1}{2}r})$, we have

$$\gamma_1 x^2 - \left(1 + \frac{5\gamma_2}{8} - \frac{35\gamma_1^2}{24}\right)x - \frac{\gamma_1}{2} + \frac{\gamma_3}{8} - \frac{35\gamma_1\gamma_2}{48} + \frac{35\gamma_1^3}{48} + O(n^{-\frac{1}{2}}) = 0.$$

Hence
$$x = -\frac{\gamma_1}{2} + \frac{\gamma_3}{8} - \frac{5\gamma_1\gamma_2}{12} + \frac{\gamma_1^3}{4} + O(n^{-2}).$$

So in the first case the mode is

$$m + \sigma\left(-\frac{\gamma_1}{2} + \frac{\gamma_3}{8} - \frac{\gamma_5}{48} + \ldots + \frac{(-)^r\gamma_{2r-1}}{2^r r!} + \ldots\right) = \kappa_1 - \frac{\kappa_3}{2\kappa_2} + \frac{\kappa_5}{8\kappa_2^3} - \frac{\kappa_7}{48\kappa_2^3} + \ldots + \frac{(-)^r\kappa_{2r+1}}{2^r r!\,\kappa_2^r} + \ldots. \qquad (1)$$

* For a comparison of this expansion with that of the Charlier Type A distribution, see Cramér (1937). The symbolic relation has also been used recently by Cornish & Fisher (1937).

If $\kappa_r = O(n)$, then the mode is

$$\begin{aligned} & m - \tfrac{1}{2}\sigma\left(\gamma_1 - \frac{\gamma_3}{4} + \frac{5\gamma_1\gamma_2}{6} - \frac{\gamma_1^3}{2}\right) + O(n^{-2}) \\ &= \kappa_1 - \frac{\kappa_3}{2\kappa_2} + \frac{\kappa_5}{8\kappa_2^2} - \frac{5\kappa_3\kappa_4}{12\kappa_2^3} + \frac{\kappa_3^3}{4\kappa_2^4} + O(n^{-2}) \\ &= \mu_1' - \frac{\mu_3}{2\mu_2} + \frac{\mu_5}{8\mu_2^2} - \frac{5\mu_3\mu_4}{12\mu_2^3} + \frac{\mu_3^3}{4\mu_2^4} + O(n^{-2}) \ldots . \end{aligned} \tag{2}$$

The median is the value of x for which $\int_{-\infty}^{x} f(u)\,du = \frac{1}{2}$, or, since

$$\frac{1}{\sqrt{(2\pi)}}\int_{-\infty}^{0} e^{-\frac{1}{2}u^2}\,du = \tfrac{1}{2},$$

the value for which

$$\frac{1}{\sqrt{(2\pi)}}\int_{0}^{x} e^{-\frac{1}{2}u^2}\,du + \frac{e^{-\frac{1}{2}x^2}}{\sqrt{(2\pi)}}\left[-\frac{\gamma_1}{6}H_2(x) + \frac{\gamma_2}{24}H_3(x) + \ldots\right] = 0.$$

So

$$e^{\frac{1}{2}x^2}\int_{0}^{x} e^{-\frac{1}{2}u^2}\,du - \frac{\gamma_1}{6}H_2(x) + \frac{\gamma_2}{24}H_3(x) - \frac{\gamma_3}{120}H_4(x) + \ldots = 0.$$

If γ_r does not fall off systematically we have

$$x = -\frac{\gamma_1}{6} + \frac{\gamma_3}{40} - \frac{\gamma_5}{336} + \ldots + \frac{(-)^r\gamma_{2r-1}}{2^r(2r+1)\,r!} + \ldots .$$

When $\kappa_r = O(n)$, we find

$$\frac{x^3}{3} - \frac{\gamma_1 x^2}{6} + \left(1 + \frac{\gamma_2}{8} - \frac{5\gamma_1^2}{4}\right)x + \frac{\gamma_1}{6} - \frac{\gamma_3}{40} + \frac{5\gamma_1\gamma_2}{48} - \frac{35\gamma_1^3}{432} + O(n^{-\frac{1}{2}}) = 0.$$

Hence

$$x = -\frac{\gamma_1}{6} + \frac{\gamma_3}{40} - \frac{\gamma_1\gamma_2}{12} + \frac{17\gamma_1^3}{324} + O(n^{-\frac{1}{2}}).$$

So if powers and products can be neglected the median is

$$\begin{aligned} & m + \sigma\left(-\frac{\gamma_1}{6} + \frac{\gamma_3}{40} - \frac{\gamma_5}{336} + \ldots\right) \\ &\qquad = \kappa_1 - \frac{\kappa_3}{6\kappa_2} + \frac{\kappa_5}{40\kappa_2^2} - \frac{\kappa_7}{336\kappa_2^3} + \ldots + \frac{(-)^r\kappa_{2r+1}}{2^r(2r+1)\,r!\,\kappa_2^r} + \ldots . \end{aligned} \tag{3}$$

And when $\kappa_r = O(n)$, the median is

$$\begin{aligned} & m - \tfrac{1}{2}\sigma\left(\frac{\gamma_1}{3} - \frac{\gamma_3}{20} + \frac{\gamma_1\gamma_2}{6} - \frac{17\gamma_1^3}{162}\right) + O(n^{-2}) \\ &\qquad = \kappa_1 - \frac{\kappa_3}{6\kappa_2} + \frac{\kappa_5}{40\kappa_2^2} - \frac{\kappa_3\kappa_4}{12\kappa_2^3} + \frac{17\kappa_3^3}{324\kappa_2^4} - O(n^{-2}) \ldots \\ &\qquad = \mu_1' - \frac{\mu_3}{6\mu_2} + \frac{\mu_5}{40\mu_2^2} - \frac{\mu_3\mu_4}{12\mu_2^3} + \frac{17\mu_3^3}{324\mu_2^4} - O(n^{-2}). \end{aligned} \tag{4}$$

Thus the distance from the mean to the mode is approximately thrice that from the mean to the median if γ_2, γ_4, etc., are small quantities (i.e. β_2 is nearly 3, etc.) and if γ_3, γ_5, etc., are small compared with γ_1. This is often the case for nearly normal distributions. Thus for the best known of Type III distributions, that of χ^2 for n degrees of freedom, $\kappa_1 = n$, $\kappa_2 = 2n$, $\kappa_r = 2^{r-1}(r-1)!\,n$. Hence from equation (2) the mode is at $n-2+O(n^{-2})$. In fact it is $n-2$ exactly. For the median, equation (4) gives $n-\frac{2}{3}+\frac{32}{405n}-O(n^{-2})$. The fraction which Pearson empirically estimated at $0{\cdot}6691+0{\cdot}0094p^{-1}$ is therefore $\frac{2}{3}+\frac{8}{405p}+O(p^{-2})$.

Consider the distribution of the estimated mean, $n^{-1}\Sigma x_r$, from a sample of n, taken from a distribution with arbitrary cumulants κ_1, κ_2, κ_3, etc. The cumulants of the distribution of the mean are κ_1, κ_2/n, κ_3/n^2, etc. Hence equations (2) and (4) hold, and the mean is κ_1, the mode is

$$\kappa_1-\frac{\kappa_3}{2n\kappa_2}+\left(\frac{\kappa_5}{8\kappa_2^2}-\frac{5\kappa_3\kappa_4}{12\kappa_2^3}+\frac{\kappa_3^3}{4\kappa_2^4}\right)n^{-2}+O(n^{-3}),$$

and the median is

$$\kappa_1-\frac{\kappa_3}{6n\kappa_2}+\left(\frac{\kappa_5}{40\kappa_2^2}-\frac{\kappa_3\kappa_4}{12\kappa_2^3}+\frac{17\kappa_3^3}{324\kappa_2^4}\right)n^{-2}+O(n^{-3}).$$

The corresponding expressions in terms of the moments may be written down. The skewness is not exactly $1/n$ of the skewness of the original curve, but nearly so if n is large.

Again consider the distribution of the estimated variance from a sample of n members from a normal distribution with unknown mean, and variance σ^2. The cumulants of the distribution of the estimated variance $\frac{1}{n-1}[\Sigma x_r^2-n^{-1}(\Sigma x_r)^2]$ are

$$\kappa_1=\sigma_1^2,\quad \kappa_2=\frac{2}{n-1},\quad \kappa_3=\frac{8\sigma^6}{(n-1)^2},\quad \kappa_4=\frac{48\sigma^8}{(n-1)^3},\quad \kappa_5=\frac{384\sigma^{10}}{(n-1)^4},\quad \text{etc.}$$

Thus we are dealing with a slightly transformed χ^2 distribution, and the mode is $\sigma^2-\frac{2\sigma^2}{n-1}$, or $\frac{(n-3)\sigma^2}{n-1}$, exactly. From equation (4) the median is

$$\sigma^2\left[1-\frac{2}{3(n-1)}+\frac{32}{405(n-1)^2}+O(n^{-3})\right],\quad \text{or}\quad \sigma^2\left[1-\frac{2}{3n}-\frac{302}{405n^2}+O(n^{-3})\right].$$

If, however, the distribution sampled is not normal, but has a finite κ_3, etc., then the first three cumulants, in terms of those of the original distribution, are

$$\kappa_2,\quad \frac{2\kappa_2^2}{n-1}+\frac{\kappa_4}{n}\quad\text{and}\quad \frac{8\kappa_2^3}{(n-1)^2}+\frac{4(n-2)\kappa_3^2}{n(n-1)^2}+\frac{12\kappa_2\kappa_4}{n(n-1)}+\frac{\kappa_6}{n^2}.$$

Fisher (1928) gives expressions for the next three cumulants, but these become very complex, $\kappa(2^6)$ having 21 terms.

It follows that the mean is κ_2, the mode

$$\kappa_2 - \left[2\kappa_2 + \frac{4\kappa_3^2 + 8\kappa_2\kappa_4 + \kappa_6}{2(2\kappa_2^2 + \kappa_4)}\right] n^{-1} + O(n^{-2}),$$

and the median

$$\kappa_2 - \left[\frac{2\kappa_2}{3} + \frac{4\kappa_3^2 + 8\kappa_2\kappa_4 + \kappa_6}{6(2\kappa_2^2 + \kappa_4)}\right] n^{-1} + O(n^{-2}).$$

The distribution of the estimate of κ_3 is symmetrical for a symmetrical distribution. In general its mean is κ_3, its mode

$$\kappa_3 - \left[45\kappa_3 + \frac{-558\kappa_3^3 + 216\kappa_2^2\kappa_5 + 108\kappa_4\kappa_5 - 9\kappa_3\kappa_6 + 27\kappa_2\kappa_7 + \kappa_9}{2(6\kappa_2^3 + 9\kappa_3^2 + 9\kappa_2\kappa_4 + \kappa_6)}\right] n^{-1} + O(n^{-2}),$$

the median differing from the mean by $\frac{1}{3}$ of the value given above.

Finally the mean estimate of κ_4 for a sample from a normal distribution is zero, its mode $-\dfrac{36\sigma^4}{n} + O(n^{-2})$, and its median $-\dfrac{12\sigma^4}{n} + O(n^{-2})$. The general expressions for the mode and median can readily be calculated from Fisher's (1928) equations.

We now pass to some cases where Pearson's trisection rule does not hold, even approximately.

Pearson's Type I and Type IV curves are asymmetrical, and have one more adjustable parameter than Type III. Thus γ_2, i.e. $\beta_2 - 3$, can vary independently of γ_1, i.e. $\beta_1^{\frac{1}{2}}$. In consequence the curves may be nearly symmetrical, but far from normal. This is so if they approximate to Type II or Type VII, respectively. In this case the even measures of divergence γ_{2r} may be much larger than the odd measures γ_{2r+1}, which tend to zero with symmetry. Thus we cannot neglect higher cumulants, or products, as in equations 1–4. For Type IV and VII curves the higher moments are infinite, so no formulae of the given type are possible. For Types I and II a formula could be given, but direct calculation is clearly preferable.

A simpler case is that of the scalene triangular distribution whose graph is obtained by joining $(-b, 0)$, $\left(0, \dfrac{2}{a+b}\right)$, and $(a, 0)$. That is to say

$$y = \frac{2(b+x)}{b(a+b)} \quad \text{for} \quad b \leqslant x \leqslant 0, \qquad \text{and} \qquad y = \frac{2(a-x)}{a(a+b)} \quad \text{for} \quad 0 \leqslant x \leqslant a.$$

Here the mean is $\frac{1}{3}(a-b)$, the mode O, and the median, if $a > b$, is

$$a - \sqrt{[\tfrac{1}{2}a(a+b)]}, \quad \text{or} \quad \frac{a-b}{4} + \frac{(a-b)^2}{32a} + \frac{(a-b)^3}{128a^2} + \ldots.$$

That is to say, the median is one-quarter, not one-third, of the distance from the mean to the mode, if $a - b$ is small compared with a. The rth moment about zero is $\dfrac{2[a^{r+1} - (-b)^{r+1}]}{r(r+1)(a+b)}$.

If $k = \dfrac{a-b}{\sqrt{(ab)}}$, then $\gamma_1 = \dfrac{\sqrt{[2k(9+2k_3^2)]}}{5(3+k^2)^{\frac{1}{2}}}\gamma_2 = -\frac{3}{5}$, $\gamma_3 = -\dfrac{6\sqrt{[2k(9+2k^2)]}}{7(3+k^2)^{\frac{1}{2}}}$, etc.

If k is small, the odd γ's are of order k, whilst the even ones approximate to those of the symmetrical distribution. Moreover $\gamma_3 = -\frac{30}{7}\gamma_1$, so the formulae 1–4 clearly do not hold.

We see then that formulae (2) and (4) are quite useful in a special case which is important in sampling theory, but have no general validity.

SUMMARY

Expressions are obtained by which the distances of the mean and median of a skew distribution can be calculated from its moments or cumulants. The series obtained may or may not converge. They give satisfactory results for Type III distributions, and for the distributions of the mean, variance, and other cumulants as estimated from samples.

REFERENCES

CORNISH, E. A. & FISHER, R. A. (1937). Moments and cumulants in the specification of distributions. *Rev. Inst. Int. Statist.* Pt. 4.

CRAMÉR, H. (1937). *Random Variables and Probability Distributions*, Cambridge Mathematical Tracts, 86–88.

FISHER, R. A. (1928). Moments and product moments of sampling distributions. *Proc. Lond. Math. Soc.* Ser. 2, **30**, 199.

PEARSON, K. (1895). Contributions to the mathematical theory of evolution. II. Skew variation in homogeneous material. *Philos. Trans.* **186**, 343.

TABLE OF PERCENTAGE POINTS OF THE *t*-DISTRIBUTION

COMPUTED BY MAXINE MERRINGTON

THE following table has been derived from Miss Thompson's Tables of Percentage Points of the Incomplete Beta Function (*Biometrika*, **32**, 168–81), by taking

$$t = \sqrt{\{\nu_2(1-x)/\nu_1 x\}}$$

for the case $\nu_1 = 1$. t is the usual "Student" ratio based on an estimate of variance having $\nu = \nu_2$ degrees of freedom. If t_P is the quantity tabled, $P/100$ is the chance that $|t| \geqslant t_P$, i.e. represents the area in the two tails of the t-distribution. The table includes certain levels for t not previously available and should be accurate to the five significant figures shown.

ν \ P	50	25	10	5	2·5	1	0·5
1	1·00000	2·4142	6·3138	12·706	25·452	63·657	127·32
2	0·81650	1·6036	2·9200	4·3027	6·2053	9·9248	14·089
3	0·76489	1·4226	2·3534	3·1825	4·1765	5·8409	7·4533
4	0·74070	1·3444	2·1318	2·7764	3·4954	4·6041	5·5976
5	0·72669	1·3009	2·0150	2·5706	3·1634	4·0321	4·7733
6	0·71756	1·2733	1·9432	2·4469	2·9687	3·7074	4·3168
7	0·71114	1·2543	1·8946	2·3646	2·8412	3·4995	4·0293
8	0·70639	1·2403	1·8595	2·3060	2·7515	3·3554	3·8325
9	0·70272	1·2297	1·8331	2·2622	2·6850	3·2498	3·6897
10	0·69981	1·2213	1·8125	2·2281	2·6338	3·1693	3·5814
11	0·69745	1·2145	1·7959	2·2010	2·5931	3·1058	3·4966
12	0·69548	1·2089	1·7823	2·1788	2·5600	3·0545	3·4284
13	0·69384	1·2041	1·7709	2·1604	2·5326	3·0123	3·3725
14	0·69242	1·2001	1·7613	2·1448	2·5096	2·9768	3·3257
15	0·69120	1·1967	1·7530	2·1315	2·4899	2·9467	3·2860
16	0·69013	1·1937	1·7459	2·1199	2·4729	2·9208	3·2520
17	0·68919	1·1910	1·7396	2·1098	2·4581	2·8982	3·2225
18	0·68837	1·1887	1·7341	2·1009	2·4450	2·8784	3·1966
19	0·68763	1·1866	1·7291	2·0930	2·4334	2·8609	3·1737
20	0·68696	1·1848	1·7247	2·0860	2·4231	2·8453	3·1534
21	0·68635	1·1831	1·7207	2·0796	2·4138	2·8314	3·1352
22	0·68580	1·1816	1·7171	2·0739	2·4055	2·8188	3·1188
23	0·68531	1·1802	1·7139	2·0687	2·3979	2·8073	3·1040
24	0·68485	1·1789	1·7109	2·0639	2·3910	2·7969	3·0905
25	0·68443	1·1777	1·7081	2·0595	2·3846	2·7874	3·0782
26	0·68405	1·1766	1·7056	2·0555	2·3788	2·7787	3·0669
27	0·68370	1·1757	1·7033	2·0518	2·3734	2·7707	3·0565
28	0·68335	1·1748	1·7011	2·0484	2·3685	2·7633	3·0469
29	0·68304	1·1739	1·6991	2·0452	2·3638	2·7564	3·0380
30	0·68276	1·1731	1·6973	2·0423	2·3596	2·7500	3·0298
40	0·68066	1·1673	1·6839	2·0211	2·3289	2·7045	2·9712
60	0·67862	1·1616	1·6707	2·0003	2·2991	2·6603	2·9146
120	0·67656	1·1559	1·6577	1·9799	2·2699	2·6174	2·8599
∞	0·67449	1·1503	1·6449	1·9600	2·2414	2·5758	2·8070

THE PROBABILITY INTEGRAL OF THE RANGE IN SAMPLES OF n OBSERVATIONS FROM A NORMAL POPULATION

I. FOREWORD AND TABLES

By E. S. PEARSON

1. *Scope of the main table*

Denote by $x_1, x_2, \ldots, x_n$ a random sample of n observations, arranged in ascending order of magnitude, drawn from a normal or Gaussian population having for probability law

$$p(x) = \frac{1}{\sqrt{(2\pi)}\sigma} \exp\left[-\tfrac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right] \tag{1}$$

where μ and σ are respectively the mean and standard deviation of the population. The range, sometimes described as the spread, in the sample is $x_n - x_1$ and we shall write the ratio of the range to the population standard deviation as

$$w = \frac{x_n - x_1}{\sigma} \tag{2}$$

No simple expression exists for the probability law $f_n(w)$ of w, but Table 1 below gives for specific values W of w computed values of the probability *integral*

$$P_n(W) = \int_0^W f_n(w)\, dw \tag{3}$$

This expression is the chance that the range in a sample of n observations is less than a given multiple of the population standard deviation. The table has been calculated for samples with n lying between 2 and 20 and for intervals of 0·05 of W. The values of the integral are given to 4-decimal place accuracy; linear interpolation is adequate except in the neighbourhood of the two quartiles in each column. The method of calculation is described below by Dr H. O. Hartley in a separate section.

2. *Auxiliary table for special uses* (Table 2)

When dealing with samples containing only a small number of observations the range or spread may often be usefully employed as a measure of dispersion in place either of the standard (root mean square) deviation or the mean deviation. For example, an estimate of the population standard deviation may be obtained by multiplying the range in a single sample or the mean range in a number of samples by the factors a_n shown in Table 2. The accuracy of this form of estimation of σ compared with that of other methods has been discussed by Davies & Pearson (1934).

In other circumstances it may be useful to plot in serial order on a control chart the values of range obtained from successive samples, e.g. when dealing with the control of quality in mass production. For this purpose it is necessary to know certain standard probability levels for w which will serve as control limits*. Twelve of these levels expressed as percentage points and obtained by interpolation in the main Table 1, are shown in Table 2. They replace approximate limits published a few years ago (Pearson, 1932). It will be found, however, that except in the case $n = 3$†, the discrepancies between the two tables are all small. As an example, the table shows that if samples of 7 observations are randomly drawn from a population with a standard deviation σ, then in the long run only 5 % of these should have a range greater than $4{\cdot}17\sigma$, while 95 % should satisfy the inequality

$$1{\cdot}25\sigma \leqslant x_n - x_1 \leqslant 4{\cdot}49\sigma$$

Probability levels for samples with $n > 12$ have not been included as the use of range for control purposes in larger samples is of doubtful value.

* For a discussion of the use of range in problems of industrial quality control see Reports issued by the British Standards Institution (Pearson, 1935, pp. 89–90; Dudding & Jennett, 1942).

† Correct values for the case $n = 3$ were given by McKay & Pearson (1933).

Table 1. *Probability integral of the range W in normal samples of size n*

W \ n	2	3	4	5	6	7	8	9	10
0·00	0·0000	0·0000							
0·05	·0282	·0007	0·0000						
0·10	·0564	·0028	·0001						
0·15	·0845	·0062	·0004	0·0000					
0·20	·1125	·0110	·0010	·0001					
0·25	0·1403	0·0171	0·0020	0·0002					
0·30	·1680	·0245	·0034	·0004	0·0000				
0·35	·1955	·0332	·0053	·0008	·0001				
0·40	·2227	·0431	·0079	·0014	·0002	0·0000			
0·45	·2497	·0543	·0111	·0022	·0004	·0001			
0·50	0·2763	0·0666	0·0152	0·0033	0·0007	0·0002	0·0000		
0·55	·3027	·0800	·0200	·0048	·0011	·0003	·0001		
0·60	·3286	·0944	·0257	·0068	·0017	·0004	·0001	0·0000	
0·65	·3542	·1099	·0323	·0092	·0025	·0007	·0002	·0001	
0·70	·3794	·1263	·0398	·0121	·0036	·0011	·0003	·0001	0·0000
0·75	0·4041	0·1436	0·0483	0·0157	0·0050	0·0016	0·0005	0·0002	0·0001
0·80	·4284	·1616	·0578	·0200	·0068	·0023	·0008	·0002	·0001
0·85	·4522	·1805	·0682	·0250	·0090	·0032	·0011	·0004	·0001
0·90	·4755	·2000	·0797	·0309	·0117	·0044	·0016	·0006	·0002
0·95	·4983	·2201	·0922	·0375	·0150	·0059	·0023	·0009	·0003
1·00	0·5205	0·2407	0·1057	0·0450	0·0188	0·0078	0·0032	0·0013	0·0005
1·05	·5422	·2618	·1201	·0535	·0234	·0101	·0043	·0018	·0008
1·10	·5633	·2833	·1355	·0629	·0287	·0129	·0057	·0025	·0011
1·15	·5839	·3051	·1517	·0733	·0348	·0163	·0075	·0035	·0016
1·20	·6039	·3272	·1688	·0847	·0417	·0203	·0098	·0047	·0022
1·25	0·6232	0·3495	0·1868	0·0970	0·0495	0·0250	0·0125	0·0062	0·0030
1·30	·6420	·3719	·2054	·1104	·0583	·0304	·0157	·0080	·0041
1·35	·6602	·3943	·2248	·1247	·0680	·0366	·0195	·0103	·0054
1·40	·6778	·4168	·2448	·1400	·0787	·0437	·0240	·0131	·0071
1·45	·6948	·4392	·2654	·1562	·0904	·0517	·0292	·0164	·0092
1·50	0·7112	0·4614	0·2865	0·1733	0·1031	0·0606	0·0353	0·0204	0·0117
1·55	·7269	·4835	·3080	·1913	·1168	·0705	·0422	·0250	·0148
1·60	·7421	·5053	·3299	·2101	·1316	·0814	·0499	·0304	·0184
1·65	·7567	·5269	·3521	·2296	·1473	·0934	·0587	·0366	·0227
1·70	·7707	·5481	·3745	·2498	·1639	·1064	·0684	·0437	·0277
1·75	0·7841	0·5690	0·3971	0·2706	0·1815	0·1204	0·0792	0·0517	0·0336
1·80	·7969	·5894	·4197	·2920	·2000	·1355	·0910	·0607	·0403
1·85	·8092	·6094	·4423	·3138	·2193	·1516	·1039	·0707	·0479
1·90	·8209	·6290	·4649	·3361	·2394	·1686	·1178	·0818	·0565
1·95	·8321	·6480	·4874	·3587	·2602	·1867	·1329	·0940	·0661
2·00	0·8427	0·6665	0·5096	0·3816	0·2816	0·2056	0·1489	0·1072	0·0768
2·05	·8528	·6845	·5317	·4046	·3035	·2254	·1661	·1216	·0886
2·10	·8624	·7019	·5534	·4277	·3260	·2460	·1842	·1371	·1015
2·15	·8716	·7187	·5748	·4508	·3489	·2673	·2033	·1536	·1156
2·20	·8802	·7349	·5957	·4739	·3720	·2893	·2232	·1712	·1307
2·25	0·8884	0·7505	0·6163	0·4969	0·3955	0·3118	0·2440	0·1899	0·1470
2·30	·8961	·7655	·6363	·5196	·4190	·3348	·2656	·2095	·1645
2·35	·9034	·7799	·6558	·5421	·4427	·3582	·2878	·2300	·1830
2·40	·9103	·7937	·6748	·5643	·4663	·3820	·3107	·2514	·2025
2·45	·9168	·8069	·6932	·5861	·4899	·4059	·3341	·2735	·2230
2·50	0·9229	0·8195	0·7110	0·6075	0·5132	0·4300	0·3579	0·2964	0·2443

Table 1 (cont.). *Probability integral of the range W in normal samples of size n*

W \ n	11	12	13	14	15	16	17	18	19	20
0·85	0·0000									
0·90	·0001									
0·95	·0001	0·0000								
1·00	0·0002	0·0001	0·0000							
1·05	·0003	·0001	·0001							
1·10	·0005	·0002	·0001	0·0000						
1·15	·0007	·0003	·0001	·0001						
1·20	·0010	·0005	·0002	·0001	0·0000					
1·25	0·0015	0·0007	0·0004	0·0002	0·0001	0·0000				
1·30	·0021	·0010	·0005	·0003	·0001	·0001	0·0000			
1·35	·0028	·0015	·0008	·0004	·0002	·0001	·0001	0·0000		
1·40	·0038	·0021	·0011	·0006	·0003	·0002	·0001	·0001		
1·45	·0051	·0028	·0016	·0009	·0005	·0003	·0001	·0001	0·0000	
1·50	0·0067	0·0038	0·0022	0·0012	0·0007	0·0004	0·0002	0·0001	0·0001	0·0000
1·55	·0087	·0051	·0030	·0017	·0010	·0006	·0003	·0002	·0001	·0001
1·60	·0111	·0067	·0040	·0024	·0014	·0008	·0005	·0003	·0002	·0001
1·65	·0140	·0086	·0053	·0032	·0020	·0012	·0007	·0004	·0003	·0002
1·70	·0175	·0111	·0069	·0043	·0027	·0017	·0011	·0007	·0004	·0003
1·75	0·0217	0·0140	0·0090	0·0058	0·0037	0·0024	0·0015	0·0010	0·0006	0·0004
1·80	·0266	·0175	·0115	·0075	·0049	·0032	·0021	·0013	·0009	·0006
1·85	·0323	·0217	·0145	·0097	·0065	·0043	·0028	·0019	·0012	·0008
1·90	·0388	·0266	·0182	·0124	·0084	·0057	·0039	·0026	·0018	·0012
1·95	·0463	·0323	·0225	·0156	·0108	·0075	·0052	·0035	·0024	·0017
2·00	0·0548	0·0389	0·0276	0·0195	0·0137	0·0097	0·0068	0·0048	0·0033	0·0023
2·05	·0643	·0465	·0335	·0241	·0173	·0124	·0088	·0063	·0045	·0032
2·10	·0749	·0550	·0403	·0295	·0215	·0156	·0114	·0082	·0060	·0043
2·15	·0866	·0646	·0481	·0357	·0264	·0196	·0144	·0106	·0078	·0057
2·20	·0994	·0753	·0569	·0429	·0323	·0242	·0181	·0136	·0102	·0076
2·25	0·1134	0·0872	0·0669	0·0511	0·0390	0·0297	0·0226	0·0172	0·0130	0·0099
2·30	·1286	·1003	·0779	·0605	·0468	·0361	·0279	·0215	·0165	·0127
2·35	·1450	·1145	·0902	·0709	·0556	·0435	·0340	·0265	·0207	·0161
2·40	·1625	·1300	·1037	·0825	·0655	·0519	·0411	·0325	·0256	·0202
2·45	·1811	·1466	·1183	·0953	·0766	·0615	·0493	·0394	·0315	·0251
2·50	0·2007	0·1644	0·1342	0·1094	0·0890	0·0722	0·0586	0·0474	0·0383	0·0309

Table 1 (cont.). *Probability integral of the range W in normal samples of size n*

W \ n	2	3	4	5	6	7	8	9	10
2·50	0·9229	0·8195	0·7110	0·6075	0·5132	0·4300	0·3579	0·2964	0·2443
2·55	·9286	·8315	·7282	·6283	·5364	·4541	·3820	·3198	·2665
2·60	·9340	·8429	·7448	·6487	·5592	·4782	·4064	·3437	·2894
2·65	·9390	·8537	·7607	·6685	·5816	·5022	·4309	·3680	·3130
2·70	·9438	·8640	·7759	·6877	·6036	·5259	·4555	·3927	·3372
2·75	0·9482	0·8737	0·7905	0·7063	0·6252	0·5494	0·4801	0·4175	0·3617
2·80	·9523	·8828	·8045	·7242	·6461	·5725	·5044	·4425	·3867
2·85	·9561	·8915	·8177	·7415	·6665	·5952	·5286	·4675	·4119
2·90	·9597	·8996	·8304	·7580	·6863	·6174	·5525	·4923	·4372
2·95	·9630	·9073	·8424	·7739	·7055	·6390	·5760	·5171	·4625
3·00	0·9661	0·9145	0·8537	0·7891	0·7239	0·6601	0·5991	0·5415	0·4878
3·05	·9690	·9212	·8645	·8036	·7416	·6806	·6216	·5656	·5129
3·10	·9716	·9275	·8746	·8174	·7587	·7003	·6436	·5892	·5378
3·15	·9741	·9334	·8842	·8305	·7750	·7194	·6649	·6124	·5623
3·20	·9763	·9388	·8931	·8429	·7905	·7377	·6856	·6350	·5864
3·25	0·9784	0·9439	0·9016	0·8546	0·8053	0·7553	0·7055	0·6569	0·6099
3·30	·9804	·9487	·9095	·8657	·8194	·7721	·7248	·6782	·6329
3·35	·9822	·9531	·9168	·8761	·8327	·7881	·7432	·6988	·6553
3·40	·9838	·9572	·9237	·8859	·8454	·8034	·7609	·7186	·6769
3·45	·9853	·9609	·9302	·8951	·8573	·8179	·7778	·7376	·6978
3·50	0·9867	0·9644	0·9361	0·9037	0·8685	0·8316	0·7939	0·7558	0·7180
3·55	·9879	·9677	·9417	·9117	·8790	·8446	·8091	·7732	·7373
3·60	·9891	·9706	·9468	·9192	·8889	·8568	·8236	·7898	·7558
3·65	·9901	·9734	·9516	·9261	·8981	·8683	·8372	·8055	·7735
3·70	·9911	·9759	·9559	·9326	·9067	·8790	·8501	·8204	·7902
3·75	0·9920	0·9782	0·9600	0·9386	0·9148	0·8891	0·8622	0·8345	0·8062
3·80	·9928	·9803	·9637	·9441	·9222	·8985	·8736	·8477	·8212
3·85	·9935	·9822	·9672	·9493	·9291	·9073	·8842	·8602	·8355
3·90	·9942	·9839	·9703	·9540	·9355	·9155	·8941	·8718	·8488
3·95	·9948	·9856	·9732	·9583	·9415	·9230	·9034	·8827	·8614
4·00	0·9953	0·9870	0·9758	0·9623	0·9469	0·9300	0·9120	0·8929	0·8731
4·05	·9958	·9883	·9782	·9660	·9519	·9365	·9199	·9024	·8841
4·10	·9963	·9895	·9804	·9693	·9566	·9425	·9273	·9112	·8943
4·15	·9967	·9906	·9824	·9724	·9608	·9480	·9341	·9193	·9038
4·20	·9970	·9916	·9842	·9752	·9647	·9530	·9404	·9269	·9126
4·25	0·9974	0·9925	0·9859	0·9777	0·9682	0·9576	0·9461	0·9338	0·9208
4·30	·9976	·9933	·9874	·9800	·9715	·9619	·9514	·9402	·9283
4·35	·9979	·9941	·9887	·9821	·9744	·9657	·9562	·9460	·9352
4·40	·9981	·9947	·9899	·9840	·9771	·9692	·9607	·9514	·9416
4·45	·9984	·9953	·9910	·9857	·9795	·9724	·9647	·9563	·9474
4·50	0·9985	0·9958	0·9920	0·9873	0·9817	0·9754	0·9684	0·9608	0·9527
4·55	·9987	·9963	·9929	·9887	·9837	·9780	·9717	·9649	·9575
4·60	·9989	·9967	·9937	·9899	·9855	·9804	·9747	·9686	·9620
4·65	·9990	·9971	·9944	·9911	·9871	·9825	·9775	·9719	·9660
4·70	·9991	·9974	·9951	·9921	·9885	·9845	·9799	·9750	·9696
4·75	0·9992	0·9977	0·9956	0·9930	0·9898	0·9862	0·9822	0·9777	0·9729
4·80	·9993	·9980	·9962	·9938	·9910	·9878	·9842	·9802	·9759
4·85	·9994	·9983	·9966	·9945	·9920	·9892	·9860	·9824	·9786
4·90	·9995	·9985	·9970	·9952	·9930	·9904	·9876	·9844	·9810
4·95	·9995	·9987	·9974	·9958	·9938	·9916	·9890	·9862	·9832
5·00	0·9996	0·9988	0·9977	0·9963	0·9946	0·9926	0·9903	0·9878	0·9851

Table 1 (cont.). *Probability integral of the range W in normal samples of size n*

W \ n	11	12	13	14	15	16	17	18	19	20
2·50	0·2007	0·1644	0·1342	0·1094	0·0890	0·0722	0·0586	0·0474	0·0383	0·0309
2·55	·2213	·1833	·1514	·1247	·1026	·0842	·0690	·0565	·0462	·0377
2·60	·2429	·2033	·1697	·1413	·1174	·0974	·0807	·0668	·0552	·0455
2·65	·2653	·2243	·1891	·1591	·1336	·1120	·0937	·0783	·0654	·0545
2·70	·2885	·2462	·2096	·1780	·1509	·1278	·1080	·0911	·0768	·0647
2·75	0·3124	0·2690	0·2311	0·1981	0·1696	0·1449	0·1236	0·1053	0·0896	0·0761
2·80	·3368	·2926	·2536	·2194	·1894	·1632	·1405	·1208	·1037	·0889
2·85	·3617	·3169	·2770	·2416	·2103	·1829	·1587	·1376	·1192	·1031
2·90	·3870	·3417	·3011	·2647	·2324	·2036	·1782	·1558	·1360	·1186
2·95	·4126	·3670	·3258	·2887	·2554	·2255	·1989	·1752	·1542	·1355
3·00	0·4382	0·3927	0·3512	0·3134	0·2792	0·2484	0·2207	0·1959	0·1737	0·1538
3·05	·4639	·4186	·3769	·3387	·3039	·2723	·2436	·2178	·1944	·1734
3·10	·4895	·4446	·4029	·3645	·3292	·2970	·2675	·2407	·2164	·1943
3·15	·5150	·4706	·4292	·3907	·3551	·3224	·2923	·2647	·2394	·2164
3·20	·5401	·4965	·4555	·4171	·3814	·3483	·3177	·2895	·2635	·2396
3·25	0·5649	0·5222	0·4817	0·4437	0·4081	0·3748	0·3438	0·3151	0·2885	0·2638
3·30	·5893	·5475	·5078	·4703	·4348	·4016	·3704	·3413	·3142	·2890
3·35	·6131	·5725	·5337	·4967	·4617	·4286	·3974	·3681	·3407	·3150
3·40	·6363	·5970	·5592	·5230	·4885	·4557	·4246	·3953	·3677	·3417
3·45	·6589	·6209	·5842	·5489	·5151	·4827	·4519	·4227	·3950	·3689
3·50	0·6807	0·6442	0·6087	0·5744	0·5413	0·5096	0·4792	0·4502	0·4226	0·3964
3·55	·7017	·6668	·6326	·5994	·5672	·5362	·5063	·4777	·4504	·4242
3·60	·7220	·6886	·6558	·6237	·5926	·5624	·5332	·5051	·4781	·4522
3·65	·7414	·7096	·6782	·6474	·6173	·5881	·5596	·5321	·5056	·4801
3·70	·7600	·7298	·6998	·6704	·6414	·6132	·5856	·5588	·5329	·5078
3·75	0·7776	0·7491	0·7206	0·6925	0·6648	0·6376	0·6110	0·5850	0·5598	0·5352
3·80	·7944	·7675	·7406	·7138	·6873	·6613	·6357	·6106	·5861	·5622
3·85	·8103	·7850	·7596	·7342	·7090	·6841	·6596	·6355	·6118	·5887
3·90	·8254	·8016	·7777	·7537	·7298	·7061	·6827	·6596	·6369	·6145
3·95	·8395	·8173	·7948	·7723	·7497	·7273	·7050	·6829	·6611	·6397
4·00	0·8528	0·8321	0·8111	0·7899	0·7686	0·7474	0·7263	0·7053	0·6845	0·6640
4·05	·8653	·8460	·8264	·8065	·7866	·7666	·7466	·7268	·7070	·6874
4·10	·8769	·8590	·8408	·8223	·8036	·7848	·7660	·7472	·7285	·7099
4·15	·8878	·8712	·8543	·8371	·8196	·8021	·7844	·7667	·7491	·7315
4·20	·8978	·8826	·8669	·8509	·8347	·8183	·8018	·7852	·7686	·7520
4·25	0·9072	0·8931	0·8787	0·8639	0·8488	0·8336	0·8182	0·8027	0·7871	0·7715
4·30	·9159	·9029	·8896	·8760	·8620	·8479	·8336	·8191	·8046	·7899
4·35	·9238	·9120	·8998	·8872	·8744	·8613	·8480	·8345	·8210	·8073
4·40	·9312	·9204	·9092	·8976	·8858	·8737	·8614	·8490	·8364	·8237
4·45	·9379	·9281	·9178	·9073	·8964	·8853	·8740	·8625	·8508	·8391
4·50	0·9441	0·9352	0·9258	0·9162	0·9062	0·8960	0·8856	0·8750	0·8643	0·8534
4·55	·9498	·9417	·9332	·9244	·9153	·9060	·8964	·8867	·8768	·8667
4·60	·9550	·9476	·9399	·9319	·9236	·9151	·9064	·8975	·8884	·8791
4·65	·9597	·9530	·9460	·9388	·9313	·9235	·9155	·9074	·8991	·8906
4·70	·9640	·9579	·9516	·9451	·9383	·9312	·9240	·9165	·9090	·9012
4·75	0·9678	0·9624	0·9567	0·9508	0·9446	0·9383	0·9317	0·9249	0·9180	0·9110
4·80	·9713	·9665	·9614	·9560	·9505	·9447	·9387	·9326	·9264	·9199
4·85	·9745	·9702	·9656	·9608	·9558	·9505	·9452	·9396	·9340	·9281
4·90	·9774	·9735	·9694	·9650	·9605	·9559	·9510	·9460	·9409	·9356
4·95	·9799	·9765	·9728	·9689	·9649	·9607	·9563	·9518	·9472	·9424
5·00	0·9822	0·9791	0·9759	0·9724	0·9688	0·9650	0·9611	0·9571	0·9529	0·9486

Table 1 (cont.). *Probability integral of the range W in normal samples of size n*

W \ n	2	3	4	5	6	7	8	9	10
5·00	0·9996	0·9988	0·9977	0·9963	0·9946	0·9926	0·9903	0·9878	0·9851
5·05	·9996	·9990	·9980	·9967	·9952	·9935	·9915	·9893	·9869
5·10	·9997	·9991	·9982	·9971	·9958	·9942	·9925	·9906	·9884
5·15	·9997	·9992	·9985	·9975	·9963	·9950	·9934	·9917	·9898
5·20	·9998	·9993	·9986	·9978	·9968	·9956	·9942	·9927	·9911
5·25	0·9998	0·9994	0·9988	0·9981	0·9972	0·9961	0·9949	0·9936	0·9922
5·30	·9998	·9995	·9990	·9983	·9975	·9966	·9956	·9944	·9931
5·35	·9998	·9995	·9991	·9985	·9979	·9971	·9961	·9951	·9940
5·40	·9999	·9996	·9992	·9987	·9981	·9974	·9966	·9957	·9948
5·45	·9999	·9997	·9993	·9989	·9984	·9978	·9971	·9963	·9954
5·50	0·9999	0·9997	0·9994	0·9991	0·9986	0·9981	0·9975	0·9968	0·9960
5·55	·9999	·9997	·9995	·9992	·9988	·9983	·9978	·9972	·9965
5·60	·9999	·9998	·9996	·9993	·9989	·9985	·9981	·9976	·9970
5·65	·9999	·9998	·9996	·9994	·9991	·9987	·9983	·9979	·9974
5·70	0·9999	·9998	·9997	·9995	·9992	·9989	·9986	·9982	·9977
5·75	1·0000	0·9999	0·9997	0·9995	0·9993	0·9991	0·9988	0·9984	0·9981
5·80		·9999	·9998	·9996	·9994	·9992	·9989	·9986	·9983
5·85		·9999	·9998	·9997	·9995	·9993	·9991	·9988	·9986
5·90		·9999	·9998	·9997	·9996	·9994	·9992	·9990	·9988
5·95		·9999	·9998	·9998	·9996	·9995	·9993	·9991	·9989
6·00		0·9999	0·9999	0·9998	0·9997	0·9996	0·9994	0·9993	0·9991
6·05		·9999	·9999	·9998	·9997	·9996	·9995	·9994	·9992
6·10		0·9999	·9999	·9998	·9998	·9997	·9996	·9995	·9993
6·15		1·0000	·9999	·9999	·9998	·9997	·9996	·9995	·9994
6·20			·9999	·9999	·9998	·9998	·9997	·9996	·9995
6·25			0·9999	0·9999	0·9999	0·9998	0·9997	0·9997	0·9996
6·30			0·9999	·9999	·9999	·9998	·9998	·9997	·9996
6·35			1·0000	·9999	·9999	·9999	·9998	·9998	·9997
6·40				0·9999	·9999	·9999	·9998	·9998	·9997
6·45				1·0000	·9999	·9999	·9999	·9998	·9998
6·50					0·9999	0·9999	0·9999	0·9999	0·9998
6·55					0·9999	·9999	·9999	·9999	·9998
6·60					1·0000	·9999	·9999	·9999	·9999
6·65						0·9999	·9999	·9999	·9999
6·70						1·0000	0·9999	·9999	·9999
6·75							1·0000	0·9999	0·9999
6·80								0·9999	·9999
6·85								1·0000	·9999
6·90									0·9999
6·95									1·0000
7·00									
7·05									
7·10									
7·15									
7·20									
7·25									

Table 1 (cont.). *Probability integral of the range W in normal samples of size n*

W \ n	11	12	13	14	15	16	17	18	19	20
5·00	0·9822	0·9791	0·9759	0·9724	0·9688	0·9650	0·9611	0·9571	0·9529	0·9486
5·05	·9843	·9815	·9786	·9756	·9723	·9690	·9655	·9618	·9581	·9543
5·10	·9861	·9837	·9811	·9784	·9755	·9725	·9694	·9661	·9628	·9593
5·15	·9878	·9856	·9833	·9809	·9783	·9757	·9729	·9700	·9670	·9639
5·20	·9893	·9874	·9853	·9832	·9809	·9785	·9760	·9735	·9708	·9681
5·25	0·9906	0·9889	0·9871	0·9852	0·9832	0·9811	0·9789	0·9766	0·9742	0·9718
5·30	·9917	·9903	·9887	·9870	·9852	·9833	·9814	·9794	·9773	·9751
5·35	·9928	·9915	·9901	·9886	·9870	·9854	·9836	·9819	·9800	·9781
5·40	·9937	·9925	·9913	·9900	·9886	·9872	·9856	·9841	·9824	·9807
5·45	·9945	·9935	·9924	·9912	·9900	·9888	·9874	·9860	·9846	·9831
5·50	0·9952	0·9943	0·9934	0·9924	0·9913	0·9902	0·9890	0·9878	0·9865	0·9852
5·55	·9958	·9951	·9942	·9933	·9924	·9914	·9904	·9893	·9882	·9870
5·60	·9964	·9957	·9950	·9942	·9934	·9925	·9916	·9907	·9897	·9887
5·65	·9969	·9963	·9956	·9950	·9943	·9935	·9927	·9919	·9910	·9901
5·70	·9973	·9968	·9962	·9956	·9950	·9944	·9937	·9929	·9922	·9914
5·75	0·9976	0·9972	0·9967	0·9962	0·9957	0·9951	0·9945	0·9939	0·9932	0·9925
5·80	·9980	·9976	·9972	·9967	·9963	·9958	·9952	·9947	·9941	·9935
5·85	·9982	·9979	·9976	·9972	·9968	·9963	·9959	·9954	·9949	·9944
5·90	·9985	·9982	·9979	·9976	·9972	·9968	·9964	·9960	·9956	·9952
5·95	·9987	·9985	·9982	·9979	·9976	·9973	·9969	·9966	·9962	·9958
6·00	0·9989	0·9987	0·9984	0·9982	0·9979	0·9977	0·9974	0·9971	0·9967	0·9964
6·05	·9990	·9989	·9987	·9984	·9982	·9980	·9977	·9975	·9972	·9969
6·10	·9992	·9990	·9989	·9987	·9985	·9983	·9981	·9978	·9976	·9973
6·15	·9993	·9992	·9990	·9989	·9987	·9985	·9983	·9981	·9979	·9977
6·20	·9994	·9993	·9992	·9990	·9989	·9987	·9986	·9984	·9982	·9980
6·25	0·9995	0·9994	0·9993	0·9992	0·9991	0·9989	0·9988	0·9986	0·9985	0·9983
6·30	·9996	·9995	·9994	·9993	·9992	·9991	·9990	·9988	·9987	·9986
6·35	·9996	·9996	·9995	·9994	·9993	·9992	·9991	·9990	·9989	·9988
6·40	·9997	·9996	·9996	·9995	·9994	·9993	·9992	·9992	·9991	·9990
6·45	·9997	·9997	·9996	·9996	·9995	·9994	·9994	·9993	·9992	·9991
6·50	0·9998	0·9997	0·9997	0·9996	0·9996	0·9995	0·9995	0·9994	0·9993	0·9993
6·55	·9998	·9998	·9997	·9997	·9996	·9996	·9995	·9995	·9994	·9994
6·60	·9998	·9998	·9998	·9997	·9997	·9997	·9996	·9996	·9995	·9995
6·65	·9999	·9998	·9998	·9998	·9997	·9997	·9997	·9996	·9996	·9995
6·70	·9999	·9999	·9998	·9998	·9998	·9998	·9997	·9997	·9997	·9996
6·75	0·9999	0·9999	0·9999	0·9999	0·9998	0·9998	0·9998	0·9997	0·9997	0·9997
6·80	·9999	·9999	·9999	·9999	·9998	·9998	·9998	·9998	·9998	·9997
6·85	·9999	·9999	·9999	·9999	·9999	·9999	·9998	·9998	·9998	·9998
6·90	0·9999	·9999	·9999	·9999	·9999	·9999	·9999	·9998	·9998	·9998
6·95	1·0000	0·9999	·9999	·9999	·9999	·9999	·9999	·9999	·9999	·9998
7·00		1·0000	0·9999	0·9999	0·9999	0·9999	0·9999	0·9999	0·9999	0·9999
7·05			0·9999	0·9999	·9999	·9999	·9999	·9999	·9999	·9999
7·10			1·0000	1·0000	0·9999	0·9999	·9999	·9999	·9999	·9999
7·15					1·0000	1·0000	0·9999	0·9999	·9999	·9999
7·20							1·0000	1·0000	0·9999	0·9999
7·25									1·0000	1·0000

3. *The origin of the present tables*

Tables giving the expected or mean value and the standard deviation of range in random samples from the normal population of equation (1) were calculated by L. H. C. Tippett (1925) in the Department of Applied Statistics, University College, London. Since the probability distribution $f_n(w)$ is itself far from normal in form, it was evident that its mean and standard deviation alone would not provide all the information generally needed in practice. Tippett included in his paper some values of the constants β_1 and β_2 of the distribution and his work was extended by the present writer (Pearson, 1926, 1932) who

Table 2

Size of sample n	Factor a_n	Lower percentage points						Upper percentage points					
		0·1	0·5	1·0	2·5	5·0	10·0	10·0	5·0	2·5	1·0	0·5	0·1
2	0·8862	0·00	0·01	0·02	0·04	0·09	0·18	2·33	2·77	3·17	3·64	3·97	4·65
3	0·5908	0·06	0·13	0·19	0·30	0·43	0·62	2·90	3·31	3·68	4·12	4·42	5·06
4	0·4857	0·20	0·34	0·43	0·59	0·76	0·98	3·24	3·63	3·98	4·40	4·69	5·31
5	0·4299	0·37	0·55	0·66	0·85	1·03	1·26	3·48	3·86	4·20	4·60	4·89	5·48
6	0·3946	0·54	0·75	0·87	1·06	1·25	1·49	3·66	4·03	4·36	4·76	5·03	5·62
7	0·3698	0·69	0·92	1·05	1·25	1·44	1·68	3·81	4·17	4·49	4·88	5·15	5·73
8	0·3512	0·83	1·08	1·20	1·41	1·60	1·83	3·93	4·29	4·61	4·99	5·26	5·82
9	0·3367	0·96	1·21	1·34	1·55	1·74	1·97	4·04	4·39	4·70	5·08	5·34	5·90
10	0·3249	1·08	1·33	1·47	1·67	1·86	2·09	4·13	4·47	4·79	5·16	5·42	5·97
11	0·3152	1·20	1·45	1·58	1·78	1·97	2·20	4·21	4·55	4·86	5·23	5·49	6·04
12	0·3069	1·30	1·55	1·68	1·88	2·07	2·30	4·29	4·62	4·92	5·29	5·54	6·09

Estimate of $\sigma = a_n \times$ range (or mean range) in a sample of n observations.

developed an approximate method of determining probability levels for w and provided some provisional tables of these. The need has, however, been felt for some time for a full and accurate table of the probability integral of the range to fit into place among other fundamental tables associated with the normal distribution. The completion of this objective has been made possible by a grant from the Department of Scientific and Industrial Research, whose assistance in the matter is acknowledged with warm appreciation. The actual method of computation was planned by Dr H. O. Hartley and the calculations were carried out under his supervision by Scientific Computing Service, Ltd. The scope of the main table was limited to $n \leqslant 20$. As n increases beyond this value there is an increasing risk that the table may be misleading in practice, since $f_n(w)$ becomes very sensitive to relatively slight departures from normality in the tails of the population distribution.

REFERENCES

DAVIES, O. L. & PEARSON, E. S. (1934). *J.R. Statist. Soc.* Supp. **1**, 76.

DUDDING, B. P. & JENNETT, W. J. (1942). *The Application of Statistical Methods to Quality Control.* British Standards Institution, No. 600, (1942).

MCKAY, A. T. & PEARSON, E. S. (1933). *Biometrika*, **25**, 415.

PEARSON, E. S. (1926). *Biometrika*, **18**, 173.

—— (1932). *Biometrika*, **24**, 404.

—— (1935). *The Application of Statistical Methods to Industrial Standardization and Quality Control.* British Standards Institution, No. 600, (1935).

TIPPETT, L. H. C. (1925). *Biometrika*, **17**, 364.

II. NUMERICAL EVALUATION OF THE PROBABILITY INTEGRAL

By H. O. HARTLEY

The formula used for the tabulation of the probability integral $P_n(W)$ of the range in normal samples of size n is given in the paper printed on pp. 334–48 below, where it proved that

$$P_n(W) = \left(\int_{-\frac{1}{2}W}^{+\frac{1}{2}W} z(x)\,dx\right)^n + 2n\int_{\frac{1}{2}W}^{\infty} z(u)\left(\int_{u-W}^{u} z(x)\,dx\right)^{n-1} du, \tag{1}$$

where

$$z(x) = (2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}x^2}.$$

Certain properties of this formula and the facilities provided by certain modern calculating machines make this integral amenable to tabulation.

The main work consists in the evaluation of the integral

$$2n\int_{\frac{1}{2}W}^{\infty} z(u)\left(\int_{u-W}^{u} z(x)\,dx\right)^{n-1} du, \tag{2}$$

by quadrature for a two variable network of values of n and W. The range of integration is from $\frac{1}{2}W$ up to a point where the integrand

$$z(u)\left(\int_{u-W}^{u} z(x)\,dx\right)^{n-1} \tag{3}$$

vanishes to 7-decimal accuracy.* For each point of the network n, W, therefore, the integrand (3) was tabulated for a set of equidistant values of u covering the range of integration. The interval of integration was chosen as $\Delta u = 0{\cdot}2$ throughout. This was sufficient to obtain about 5-decimal accuracy in the integral (2).

The interval in W was taken as wide as possible but sufficiently fine to permit checking by differencing and the subsequent subtabulation of $P_n(W)$ to interval 0·05, which is the interval in the final table. An interval of $\Delta W = 0{\cdot}25$ was therefore chosen for the n, W network.

For small values of W it was necessary to tabulate the integrand for all integers n for which $P_n(W)$ is required in the final table. For larger values of n and W, however, it was sufficient to calculate the final integral (1) for odd n and to obtain intermediate values by interpolation. Below, then, is shown the two variable network for which the integral (2) was produced by quadrature:

$$\left.\begin{array}{ll} W = 0{\cdot}00\ (0{\cdot}25)\ 1{\cdot}25 & \text{and} \quad n = 3\ (1)\ 20. \\ W = 1{\cdot}50\ (0{\cdot}25)\ 2{\cdot}75 & \qquad\quad n = 3\ (1)\ \ 9\ (2)\ 23. \\ W = 3{\cdot}00\ (0{\cdot}25)\ 3{\cdot}25 & \qquad\quad n = 3\ (1)\ \ 5\ (2)\ 23. \\ W = 3{\cdot}50\ (0{\cdot}25)\ 8{\cdot}00 & \qquad\quad n = 3\ (2)\ 23. \end{array}\right\} \tag{4}$$

For $n = 2$ the final integral $P_2(W)$ is given directly by the normal integral and may be obtained by interpolation in Table II of *Tables for Statisticians and Biometricians*, Part I. Using the notation of that table (Sheppard's original notation) we have

$$P_2(W) = \alpha\left(\frac{W}{\sqrt{2}}\right).$$

Moreover, for purposes of interpolation, use was made of the formal relation

$$P_1(W) = 1 \quad \text{for} \quad W > 0.$$

For fixed u and W and for values of n in the arithmetic progression (4), the integrand (3) is a geometric progression with

$$z(u)\left(\int_{u-W}^{u} z(x)\,dx\right)^2$$

* The integrand was calculated to 7-decimal accuracy in order to obtain $P_n(W)$ to about 5-decimal accuracy.

as leading term and $\left(\int_{u-W}^{u} z(x)\,dx\right)$ or $\left(\int_{u-W}^{u} z(x)\,dx\right)^2$

as common ratio. This leading term as well as the common ratios were easily obtained from Table II of *Tables for Statisticians and Biometricians*, Part I and the terms of the progression were then automatically produced on a Mercedes calculating machine Model 38 M.S. and copied down in two-way tables with u as row heading, n as column heading and W as table heading. The values of the integrand were then checked by differencing column-wise and added to yield the main term of the integral (2). The correction terms which, according to Gregory's formula, convert the integrand-sum into the integral were calculated from the differences and checked by the application of Gauss' formula of integration. Finally, to obtain $P_n(W)$ the term

$$\left(\int_{-\frac{1}{2}W}^{+\frac{1}{2}W} z(x)\,dx\right)^n$$

was produced by continued multiplication on the Mercedes and added to the corresponding integral (2) to yield $P_n(W)$ for all points of the above network.

For odd values of n the integral $P_n(W)$ was then differenced W-wise on the National machine which, incidentally, produced column totals $\sum_W P_n(W)$ for these values of n. Two checks were applied at this stage. One consisted in inspecting the fourth order differences. As a second check, the mean range, $\overline{w}_n$, was calculated from the formula

$$\overline{w}_n = 8 - \int_0^8 P_n(w)\,dw,^*$$

and compared with the correct mean range given in Table XXII of *Tables for Statisticians and Biometricians*, Part II. Finally, the function $P_n(W)$ was subtabulated to interval 0·05 on the National machine by a method similar to that described in detail by L. J. Comrie (1936).

The values of $P_n(W)$ for even n were then obtained by interpolation with the help of two interpolation formulae of Lagrangian type:

$$\begin{aligned} 2048P_n(W) = &-5[P_{n-7}(W)+P_{n+7}(W)]+49[P_{n-5}(W)+P_{n+5}(W)] \\ &-245[P_{n-3}(W)+P_{n+3}(W)]+1225[P_{n-1}(W)+P_{n+1}(W)], \end{aligned} \tag{5}$$

$$\begin{aligned} 20P_n(W) = &P_{n-3}(W)+P_{n+3}(W)-6[P_{n-2}(W)+P_{n+2}(W)] \\ &+15[P_{n-1}(W)+P_{n+1}(W)]. \end{aligned} \tag{6}$$

Formula (5) yields the interpolate for even n from the given values of $P_n(W)$ at adjacent odd values of n. This formula was used throughout. In some cases, however, the resulting interpolate was accurate to about 3 places of decimals only. In such cases values of P_{n-3}, P_{n+3}, P_{n-1}, P_{n+1} accurate to 5 places of decimals and values of P_{n-2}, P_{n+2} accurate to (say) 3 places of decimals were substituted in formula (6). This yielded a 'corrected value' of $P_n(W)$. The process was then repeated for $n = n+2$ and so on until all values of $P_n(W)$ had 'settled down' for even values of n. It is easy to see that the process is convergent and that the maximum error in the interpolate is 2 units for the 5th decimal.

After completion of the interpolation n-wise, the interpolates $P_n(W)$ for even n were differenced W-wise, checked and subtabulated as for odd values of n.

* This is true provided $P_n(8)=1$ to 6-decimal place accuracy.

REFERENCE

COMRIE, L. J. (1936). *J.R. Statist. Soc.* Suppl. **3**, 87–114.

NOTES ON TESTING STATISTICAL HYPOTHESES

By E. S. PEARSON

1. In July 1939, a few weeks before the opening of the present war, a Conference on the Application of the Calculus of Probabilities was held at Geneva under the auspices of the International Institute of Intellectual Co-operation (League of Nations). At the public session at which a paper by Prof. J. Neyman was presented and also subsequently in some informal discussions, a number of questions were raised:

(*a*) In choosing a test for a statistical hypothesis, is it possible or even necessary to specify the hypotheses alternative to that tested? Why should not a test be made to depend only on the form of law associated with the hypothesis tested? For example, Newton's hypothesis of gravitation was formulated and tested without any need to define alternative laws.

(*b*) Is the method of approach to these problems advocated by Prof. Neyman and myself applicable to testing the appropriateness of probability *laws* or only to testing hypotheses regarding the numerical values of *constants* contained in these laws?

After the conclusion of the conference, I set down some Notes for a few of the statisticians who had taken part in the discussions, hoping that at leisure they might feel stimulated to define their views on the subject more precisely. But almost before the Notes were despatched, war in Europe had intervened. The only reply which I received was from Prof. Gumbel, and this, after some unavoidable delay, has now taken shape in the contribution printed on pp. 317–33 below. In publishing this, it seems useful to add my own Notes, which are given with only minor verbal alterations in the following pages. They are in part a restatement with rather different emphasis of views expressed in a paper published four years ago (Pearson, 1938).

2. With regard to one of the points raised under (*a*) above, it should be remembered that a statistical hypothesis as defined by Neyman and myself is a hypothesis concerning the probability law of random variables. The gravitational hypothesis of Newton is not a statistical hypothesis in the sense defined; statistical methods may be introduced to test the Newtonian hypothesis, however, and they will involve tests of statistical hypotheses or 'significance tests' because it will be assumed that errors of observation exist which may be regarded as random variables, probably taken to follow the normal distribution law.

For example, on the Newtonian hypothesis, the angular co-ordinates of a planet measured from the earth as origin may at certain moments be given as $\xi = \xi_t, \eta = \eta_t\ (t = 1, 2, \ldots)$. If we have a number of observations of position x_t, y_t,

subject to observational error, the statistical problem will be to test whether these are consistent with the hypothetical position values ξ_t, η_t, or whether they suggest that ξ, η have some other different values at the moments of observation. Thus the 'alternatives' that we have immediately in mind will be alternative values for ξ, η, not alternative gravitational hypotheses. If, however, some alternative law of motion were proposed, so that we could specify definite values ξ_t', η_t' alternative to the values ξ_t, η_t of the Newtonian law, then undoubtedly we could choose a statistical test which would be particularly efficient in discriminating between the two alternatives. Such a course became possible when the Einstein hypothesis was formulated and the orbit of Mercury considered. But the absence of an alternative gravitational law does not prevent us selecting a statistical test which will be (*a*) sensitive to departures in ξ, η from ξ_t, η_t, but (*b*) relatively insensitive to departures from normality in the distribution of errors. We should make this selection because, if the Newtonian law were incorrect, we believe that this would result in a change in ξ_t, η_t but not in a departure of the distribution of observational errors from the normal law.

This example, of course, concerns a statistical hypothesis regarding the values of two parameters ξ, η, not regarding the form of a probability law of random variables. The following general approach shows, however, that the principles discussed may be applied to testing hypotheses regarding probability laws.

3. Suppose that x is a continuous random variable and that H_0 is a statistical hypothesis which assumes that the elementary probability law for x is $p(x \mid H_0)$ in the interval $-\infty$ to $+\infty$. Thus

$$\int_{-\infty}^{+\infty} p(x \mid H_0)\, dx = 1. \tag{1}$$

Now write

$$y = \int_{-\infty}^{x} p(x \mid H_0)\, dx. \tag{2}$$

y will be a non-decreasing function of x having values confined to the interval $(0, 1)$. Further, the elementary probability law of y will be

$$p(y) = p(x)\Big/\frac{dy}{dx} = 1 \quad \text{for} \quad 0 \leqslant y \leqslant 1, \tag{3}$$

or all values of y between 0 and 1 are 'equally probable'.

Suppose now that we wish to use a set of n independent values $x_1, x_2, \ldots, x_n$ to test that the probability law is of the assumed form $p(x \mid H_0)$. It is clear that the hypothesis H_0 is exactly equivalent to the hypothesis, say h_0, that the n values $y_1, y_2, \ldots, y_n$ (obtained from the x's by the transformation (2)) have been sampled subject to the probability law (3). Just as the point $(x_1, x_2, \ldots, x_n)$ may be represented in an unlimited n-dimensioned space having probability density

$$p(x_1, x_2, \ldots, x_n \mid H_0) = \prod_{i=1}^{n} \{p(x_i \mid H_0)\}, \tag{4}$$

if H_0 is true, so the point $(y_0, y_1, \ldots, y_n)$ may be placed in an n-dimensioned hypercube with sides of unit length and with uniform probability density, if H_0 and therefore h_0 is true. It follows that if H_0 is what has been termed a 'simple hypothesis', i.e. specifies the form of $p(x \mid H_0)$ completely,* then the test of H_0 may always be transformed to the test of h_0. If then it were correct to say that the test of a statistical hypothesis *depends only on the form of the law specified by* H_0, it follows that for the type of situation considered the testing of a statistical hypothesis could always be reduced to the following simple problem:

To test whether a sample of n independent random variables $y_1, y_2, \ldots, y_n$ $(0 \leqslant y_i \leqslant 1)$ has been selected from the so-called rectangular distribution, i.e. the distribution for which $p(y) = 1$, $(0 \leqslant y \leqslant 1)$.

4. We are at once faced, therefore, with the question of how to test this simple but apparently fundamental hypothesis. If h_0 is true, the sample point is equally likely to fall at any point within the n-dimensioned hypercube. Thus in picking out the critical (or rejection) region in this space we can get no assistance whatsoever from the changes in probability density, as we might do in the x-space. If we wish to use a level of significance of α (say $\alpha = 0{\cdot}01$) for rejecting h_0, it is clear that an infinite number of critical regions satisfying this condition are available; it is only necessary to select a region whose content is α.

If we consider the n values of y and plot them in the interval (0, 1) as follows,

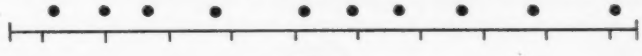


Fig. 1.

the great majority of samples, from a rectangular distribution, at any rate if n is not too small, will be spread out fairly uniformly throughout the interval. Perhaps an 'ideal' sample against which to measure irregularities might be described as one for which the values of y fell at

$$\frac{1}{2n}, \quad \frac{3}{2n}, \quad \frac{5}{2n}, \quad \ldots, \quad \frac{2n-1}{2n}.$$

But what form of departure from this ideal of uniformity are we to pick out as suggesting that the hypothesis h_0 is disproved? Should we judge significance by paying attention to the value of the mean y, of the variance, of the range of variation or of higher moments? Or should we use the χ^2 or ω^2 tests? It seems difficult to find any basis for choice which could be regarded in any sense as the 'best'. For any set of values $y_1, y_2, \ldots, y_n$ *some* critical region of size α can always be found which will contain the sample point and therefore lead to the rejection of h_0. Indeed, the task of selecting a unique region on any rational basis would seem to be insoluble.

* This condition is important. If the values of certain constants contained in the probability law need to be estimated from the observations, then the n values of y will not form a true random sample from a rectangular distribution. They will be subject to certain limitations to their degrees of freedom, though these may be relatively unimportant if n is large.

5. Directly it is recognized, however, that the choice of a test of a statistical hypothesis depends on something more than the form of the law associated with that hypothesis, it can be seen how a solution may be obtained. If we can specify a single alternative H_1 to H_0 or a class of alternatives $C(H)$, then we shall have also an alternative h, or a class $C(h)$ to h_0. Thus, if $p(x \mid H_1)$ denotes a probability law alternative to $p(x \mid H_0)$, then for y the alternative is

$$p(y \mid h_1) = p(x \mid H_1) \Big/ \frac{dy}{dx} = \frac{p(x \mid H_1)}{p(x \mid H_0)}\Big|_{x=f(y)} \quad \text{for} \quad 0 \leqslant y \leqslant 1, \tag{5}$$

where $f(y)$ means the solution of

$$y = \int_{-\infty}^{x} p(x \mid H_0)\, dx \tag{6}$$

with regard to x. For example, Fig. 2 shows three typical forms of alternative $p(y \mid h_1)$, $p(y \mid h_2)$ and $p(y \mid h_3)$ associated with alternatives $p(x \mid H_1)$, ..., etc., to $p(x \mid H_0)$.

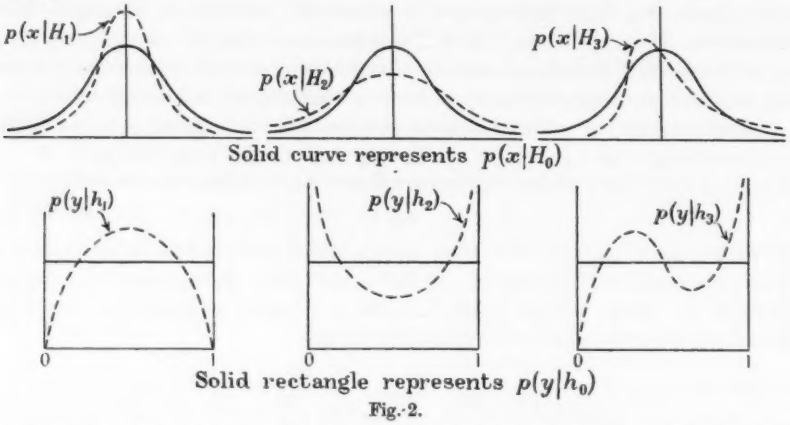


Fig. 2.

We can now see the kind of test which will be most efficient for testing H_0 with regard to possible classes of alternatives. If the alternative laws are of smaller dispersion (as $p(x \mid H_1)$), we must be on the look-out for too many values of y near $\frac{1}{2}$ and too few near 0 and 1. For alternatives with greater dispersion (as $p(x \mid H_2)$), we must reject H_0 when there are too many y's near 0 or 1 and too few near $\frac{1}{2}$. While if the alternatives are likely to be asymmetrical curves (as $p(x \mid H_3)$), then a different rule will be needed, as suggested by the $p(y \mid h_3)$ curve.

6. It follows that in so far as it is possible to formulate the class of admissible probability laws $p(x \mid H)$, the problem of selecting the most efficient test of H_0 reduces to that of choosing a critical region in the n-dimensioned hypercube which is most effective in detecting, from a sample of n values of y, differences between the rectangle $p(y \mid h_0)$ and the appropriate alternative forms $p(y \mid h)$.

If H_1 is a single admissible alternative, then it has been shown (Neyman & Pearson, 1933, p. 298) that the region w_0 of content α in the hypercube, within which

$$\frac{\prod_{i=1}^{n} p(y_i \mid h_0)}{\prod_{i=1}^{n} p(y_i \mid h_1)} < k, \tag{7}$$

or, in view of equation (3),

$$\prod_{i=1}^{n} p(y_i \mid h_1) > \frac{1}{k}, \tag{8}$$

where k is chosen so that

$$P\{(y_1, y_2, \ldots, y_n) \epsilon w_0 \mid h_0\} = \alpha \tag{9}$$

has the following property.

Of all regions of content α, w_0 is more likely than any other to include the sample point when h_1, and not h_0, is true. The region has been termed the best critical region for testing h_0 with regard to the alternative h_1.

As soon as H_0 and H_1 are specified, clearly $p(y \mid h_1)$ and therefore the region w_0 can be found, although mathematically it may be rather difficult to determine the appropriate boundary $\prod_{i=1}^{n} p(y_i \mid h_1) = \text{constant}$, so as to satisfy (8). Since this product is the probability density in the hypercube given by h_1, it will be seen that what we set out to do is to include in the critical region those parts of the sample space where the density for h_1 is highest. It is here, on repeated sampling, that sample points would tend to be concentrated if h_1 is true, instead of being uniformly distributed as under h_0.

7. If instead of a single alternative h_1, there is a class of admissible alternatives $C(h)$, there may or may not be common points of concentration that can be included in the critical region. This will depend on whether the inequality (8) above defines a region independent of the particular hypothesis h of the class $C(h)$. Even if there is no single region of content α which is exactly a 'best critical region' for h_0 with regard to all members of $C(h)$, the general principle may still be used as a guide. We build up a critical region out of those parts of the hypercube where the probability density tends to be concentrated when the probability law departs from $p(x \mid H_0)$ in the direction of the alternatives included in $C(H)$.

For example, in my earlier paper (Pearson, 1938) I suggested as appropriate in the following situation a test which, while not based on a common best critical region, was selected so as to include regions of greatest density associated with alternatives of $C(h)$. For the hypothesis tested,

$$p(x \mid H_0) = \frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}x^2}. \tag{10}$$

The alternatives are asymmetrical curves with the same mean and standard deviation as (10). A typical alternative would be the Type III curve

$$p(x \mid H) = c(1 + \tfrac{1}{2}x\sqrt{\beta_1})^{\frac{4}{\beta_1}-1} e^{-\frac{2x}{\sqrt{\beta_1}}}, \tag{11}$$

whose form departs more and more from (10) as $\sqrt{\beta_1}$ increases from zero, but the class need not be defined as precisely as this. In this problem it appears that if n independent observations $x_1, x_2, \ldots, x_n$ are available, the following is a good test of H_0. Take as test function

$$Q = \prod_{i=1}^{n} (y_i'), \tag{12}$$

where

$$\left.\begin{aligned} y_i' &= 5(0{\cdot}2 - y_i) \quad \text{for} \quad 0 \leqslant y_i \leqslant 0{\cdot}2, \\ y_i' &= \tfrac{5}{3}(y_i - 0{\cdot}2) \quad \text{for} \quad 0{\cdot}2 < y_i \leqslant 0{\cdot}8, \\ y_i' &= 5(1 - y_i) \quad \text{for} \quad 0{\cdot}8 < y_i \leqslant 1, \end{aligned}\right\} \tag{13}$$

and

$$y_i = \int_{-\infty}^{x} \frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}x^2}\, dx. \tag{14}$$

If H_0 is true it may be shown that $-2\log_e Q$ is distributed as χ^2 with $2n$ degrees of freedom. Hence any desired significance level α, for Q, may be found. We should then reject H_0 when Q is significantly small.

A more systematic method of dealing with such problems has been considered by Neyman (1937) in his paper on 'smooth tests'.

8. To sum up, the position seems to be this. It has often been argued that a statistical test need only depend on the form of the probability law associated with the hypothesis tested. In the case where H_0 concerns the probability law of a single random variable and where $p(x \mid H_0)$ is precisely specified, by the transformation from x to y it has been shown that the problem of testing H_0 on the basis of n independent values of x can always be reduced to another problem, which involves this question. Can we regard a sample $y_1, y_2, \ldots, y_n$ as having been drawn from the rectangular distribution $p(y \mid h_0) = 1$, where $0 \leqslant y \leqslant 1$? We are faced with a single fundamental question and we have to consider whether it can be answered in a rational manner, unless we are prepared to take into account the kind of departures from the rectangular law that we either believe possible or at any rate consider it most important to be on the look out for.

The transformation from x to y seems to have the advantage that it concentrates attention on the main point at issue. That is my reason for emphasizing it in these Notes. Most of us have many preconceived ideas about appropriate tests if the probability law is taken in the form of $p(x \mid H_0)$; we are accustomed to use the mean, the standard deviation, certain functions of moments, the χ^2 test, ... But we are not so accustomed to test whether a sample comes from a rectangular distribution and we are therefore forced or, indeed, more willing to reconsider from first principles what course we should follow and why.

REFERENCES

NEYMAN, J. (1937). *Skand. Actuar. Tidskr.* 149.
NEYMAN, J. & PEARSON, E. S. (1933). *Philos. Trans.* A, **231**, 289.
PEARSON, E. S. (1938). *Biometrika*, **30**, 134.

SIMPLE TESTS FOR GIVEN HYPOTHESES

By E. J. GUMBEL
New School for Social Research, New York

CONTENTS

PAGE
1. Control curves and the probability integral transformation . . . 317
2. Classical tests applied to a uniform distribution 321
3. The mth point test 326

In dealing with the problem of testing statistical hypotheses J. Neyman (1937) and E. S. Pearson (1938) have considered the use of the probability integral transformation, which leads to a theoretical uniform distribution. This method presupposes that the usual comparison between theory and observations has already been applied. We shall first improve this comparison by introducing control curves. Then we shall apply to the uniform distribution the usual methods and the control curves. This will lead to simple tests for given statistical hypotheses.

1. Control curves and the probability integral transformation

Let x be a continuous random variable for which n observations have been made. Let x_m be the observed values arranged in increasing order of magnitude, m $(1, 2, \ldots, n)$ being the serial number. The simplest way of representing the observations is to plot the cumulative histogram x_m, m. The relative number $W^{(0)}(x_m)$ of observations less than or equal to x_m is given by

$$m = nW^{(0)}(x_m).$$

The consecutive differences

$$nW^{(0)}(x_m) - nW^{(0)}(x_l) \quad (l < m)$$

constitute the observed distribution. Many statisticians present, instead of the original observations x_m, only the number of cases within certain arbitrary classes. From the practical standpoint this means a simplification, from the theoretical standpoint a complication. We shall suppose that all x_m are known.

The choice of a probability density to be applied to the observations constitutes the hypothesis. The probability density $w(x, c_1, c_2, \ldots)$, where c_1, c_2, ... are the constants, is called the theoretical distribution. For sake of simplicity it is assumed that all observed values x_m have the same theoretical distribution. The probability $W(x, c_1, c_2, \ldots)$ of a value equal to or less than x is given by

$$W(x, c_1, c_2, \ldots) = \int_{-\infty}^{x} w(z, c_1, c_2, \ldots)\, dz, \tag{1}$$

where z is the variable of integration. It is customary to compare the observations x_m, m with the cumulative frequency curve x, $nW(x)$. This comparison can be improved in the following way: The mth observation is a statistical variable distributed according to

$$\mathfrak{w}_m(x, n) = \binom{n}{m} m W(x)^{m-1}(1 - W(x))^{n-m} w(x). \tag{2}$$

In a previous article (Gumbel, 1935) it has been shown that, for an ordinary unlimited distribution, with large n and with m of the size $\frac{1}{2}n$, the distribution (2) converges towards a normal distribution with a mean given by $W(x) = m/n$ and a standard deviation

$$\sigma = \frac{1}{w(x)} \sqrt{\frac{W(x)\,(1 - W(x))}{n}}. \tag{3}$$

This formula does not contain m explicitly. Since each theoretical value x can be interpreted as an mth value, (3) gives its standard deviation. The interval $x \mp \sigma$ will be called the *control interval*. Under the above condition, the probability that an mth value will fall within the control interval is about $\frac{2}{3}$. The two curves obtained by plotting $x \mp \sigma$, $nW(x)$ will be called control curves.

For a given initial distribution we shall have to find the mean and the standard deviation of the mth value which may differ from those of the general solution, especially if n is small. The control interval will be a certain function of $w(x)$. Also the probability associated with the interval $x \mp \sigma$ may differ from that of the general solution and may depend upon x. For the exponential distribution (Gumbel, 1937) the precision diminishes with increasing values of the variable. Below we shall apply this control to the uniform distribution $w(x) = \text{constant}$.

The calculation of the probability $W(x)$ and the control curves can often be simplified by an indirect method. For certain, but not all, distributions it is possible to eliminate the constants by introducing a new variable y as a function of x, where

$$x = f(y, c_1, c_2, \ldots). \tag{4}$$

Accordingly, the probability that a value of the transformed variable will fall in the interval y to $y + dy$ is

$$w(x)\,dx = w[f(y)]\,|\,f'(y)\,|\,dy.$$

We call

$$p(y) = w[f(y)]\,|\,f'(y)\,|,$$

or

$$p(y) = w(x)\left|\frac{dx}{dy}\right|, \tag{5}$$

the distribution of y. The probability $V(y)$ of a value equal to or less than y is

$$V(y) = W(x, c_1, c_2, \ldots). \tag{6}$$

The transformation (4) is chosen such that the expression $V(y)$ does not contain any constants which depend upon the observations. Therefore $V(y)$ can be

calculated once and for all as a function of y. Such tables have been calculated for the distributions for which this reduction is possible.

In order to compare the cumulative histogram x_m, m with the cumulative frequency x, $nW(x)$, and to use the control curves, it is first necessary to compute the constants c_1, c_2, ... If the method of moments is used, the area between the cumulative histogram and the horizontal line $W = 1$, the arithmetic mean, is conserved. The value of the variable x which corresponds to a selected, numerical value of $W(x)$ is obtained from the transformation (4).

A special case of the transformation (4) which leads to a reduced distribution of astounding simplicity is the probability integral transformation due to Karl Pearson (1933) who introduced for y the probability function $W(x)$. Since

$$\frac{dx}{dW} = \frac{1}{w(x)}, \tag{7}$$

we obtain from (5)

$$p(W) \equiv 1. \tag{8}$$

This identity means that the distribution of the probability is constant. As $p(W)$ is the probability density of a probability, it is difficult to establish its philosophical meaning. But formally the construction is valid and the corresponding value can be observed. It is our purpose to give several methods of judging the significance of the differences between the theoretical distribution (8) and the corresponding 'observed' distribution.

The word 'observation' will be given a special meaning. A certain theory h_0 which involves the choice of certain constants c_1, c_2, ... applied to the observations, leads to the values

$$W_0 = W(x_m, c_1, c_2, \ldots \mid h_0) \quad (m = 1, 2, \ldots, n),$$

corresponding to x_m. These values, contained in the interval 0, 1, are the 'observations'. To any other set of constants c_1', c_2', ... will correspond other 'observed' values

$$W_0' = W(x_m, c_1', c_2', \ldots \mid h_0).$$

Therefore any test applied to the 'observations' of formula (8) might be used to judge the choice of the constants. To another hypothesis h_1 containing the constants d_1, d_2, ..., will correspond another set of 'observed' points

$$W_1 = W(x_m, d_1, d_2, \ldots \mid h_1).$$

The same observations when interpreted by different theories or different constants lead to different 'observations'. An incorrect theory involving properly chosen constants might give better results than a correct theory involving improperly chosen constants. Therefore, to compare different theories, the constants for each must be determined with the same precision. In practice this condition will not always be fulfilled. For the precision of the characteristics depends upon the distribution for which they are calculated (Gumbel, 1936). Therefore, the

same method of determining the constants might lead to different degrees of precision for different distributions, whereas different determinations of the constants might lead to approximately the same precision.

There is another point for caution: all tests derived from the probability integral transformation apply to the analytic form of the hypotheses and at the same time to the choice of the constants. A formula containing many constants may reproduce the observations more closely than a formula containing few constants, even though the constants in the first hypothesis have no meaning. Therefore we must limit our comparison to hypotheses containing the same number of constants. For any set of statistical observations in the ordinary sense, there will usually correspond a small number of tenable hypotheses. We shall suppose it is known what they are. For we do not try to find a formula for the sake of doing it, but to explain the observed facts. We will not go so far as Neyman (1937), who formulated all possible alternatives by a series of orthogonal functions.

In theory the points representing $W(x_m)$ are distributed uniformly in the interval 0, 1. This is true for any hypothesis, provided the variable is continuous. But in practice this will never occur. The 'observed' points corresponding to any given hypothesis will differ from the theoretical set, even if the hypothesis is a very good one. The differences between the 'observed' set of points resulting from h_0 and h_1 and the theoretical set allow the construction of tests which can be used to judge which of two given hypotheses is the better. But no statistical method gives an answer to the question whether or not a hypothesis is true.

After a hypothesis has been selected, the preliminary steps which have to be made before it is possible to use the probability integral transformation, are: first, the determination of the constants; secondly, the calculation of probabilities $W(x)$ for the values of x given by the transformation (4); and thirdly, the calculation of the probabilities $W(x_m)$ of the observed values. It is only after these three operations have been carried out that we obtain the 'observations' which are to be compared with the theory (8). Therefore, any test based on the probability integral transformation presupposes the usual comparison of the observed cumulative histogram with the frequency curve. In many cases this comparison, checked by the control curves, will indicate a clear superiority of one of the theories. If this is true, there is no necessity for a new test.

It would be interesting to investigate the best criterion for judging the significance of the differences between the 'observed' and the uniform distribution (8). But for practical purposes it is sufficient to know whether the differences for h_0 are smaller or larger than for h_1. First we shall establish rough measures of comparison; afterwards, more refined ones.

2. CLASSICAL TESTS APPLIED TO A UNIFORM DISTRIBUTION

The comparison of an observed distribution of a continuous variable with the theoretical distribution is reduced by the probability integral transformation to a comparison of 'observed' points with a uniform distribution. It seems logical to use first the classical methods which are here very simple, as no constants have to be determined. For a uniform distribution

$$p(y) = 1 \quad (0 \leqq y \leqq 1),$$

the arithmetic mean and the median are, respectively,

$$\bar{y} = \tilde{y} = \tfrac{1}{2}. \tag{9}$$

The mean error θ and the probable error ρ, defined as half of the difference between the two quartiles, are

$$\theta = \rho = \tfrac{1}{4}. \tag{10}$$

The kth moment about the origin is

$$M_k = \frac{1}{k+1},$$

which gives the recurrence relation

$$\frac{1}{M_{k+1}} - \frac{1}{M_k} = 1. \tag{11}$$

Since the distribution is symmetrical, the odd moments about the arithmetic mean vanish. Therefore

$$\beta_1 = 0.$$

The even moments are

$$\mu_{2k} = \int_0^1 (y - \tfrac{1}{2})^{2k}\, dy = 2\int_0^{\frac{1}{2}} z^{2k} dz \quad \text{or} \quad \mu_{2k} = \frac{1}{2^{2k}(2k+1)}. \tag{12}$$

Therefore the standard deviation, the coefficient of variation and the second beta are, respectively,

$$\sigma = \frac{1}{2\sqrt{3}}, \tag{13}$$

$$v = \frac{1}{\sqrt{3}}, \tag{14}$$

and

$$\beta_2 = \tfrac{9}{5}. \tag{15}$$

It is necessary now to calculate the 'observed' means, the measures of dispersion, and the relations between successive moments. To control the agreement between the theoretical uniform distribution and the 'observed' points we can still employ the standard error $\sigma_{\bar{y}}$ of the arithmetic mean. The general formula

$$\sigma_{\bar{y}} = \frac{\sigma}{\sqrt{n}}$$

becomes according to (13) $$\sigma_{\bar{y}} = \frac{1}{2\sqrt{(3n)}}. \tag{16}$$

The standard error of the dispersion for n large is $\sigma(\sigma^2) = \sqrt{\frac{\mu_4 - \mu_2^2}{n}}$, which becomes, according to (12),

$$\sigma(\sigma^2) = \sqrt{\left\{\frac{1}{4^2 n}\left(\frac{1}{5} - \frac{1}{9}\right)\right\}} = \frac{1}{6\sqrt{(5n)}}. \tag{17}$$

It seems reasonable to employ these old-fashioned tests before the use of more sophisticated methods is resorted to. Only if they fail would it be necessary to consider more elaborate methods.

The n 'observed' points $W(x_m)$ represent the probabilities, obtained from the hypothesis h_0, of the given observations in the ordinary sense, x_m. We plot these points in the interval 0, 1 which is divided into k cells of equal length, where k is chosen in such a way that n/k is an integer. If n is a multiple of 10, we choose the cells (0·0, 0·1), (0·1, 0·2), ..., (0·9, 1·0).

The probability density of a point falling somewhere within the interval 0, 1 is constant. As the interval is of length 1, this density is 1. Therefore, the probability of a point lying with a given cell is $1/k$, and the expected number of points in each cell is n/k. The 'observed' number of points obtained through a hypothesis h_0 will be a_ν ($\nu = 1, 2, \ldots, k$). If we apply another hypothesis h_1 to the same observations or introduce other numerical values for the constants, the new set of 'observed' points will lead to values b_ν which, in general, will differ from a_ν.

The classical statistical method of treating this material is the χ^2 test. As the numbers a_ν, b_ν will differ from the expected number n/k, we can calculate for both hypotheses

$$\chi_0^2 = \frac{k}{n}\sum_{\nu=1}^{k}\left(a_\nu - \frac{n}{k}\right)^2, \quad \chi_1^2 = \frac{k}{n}\sum_{\nu=1}^{k}\left(b_\nu - \frac{n}{k}\right)^2. \tag{18}$$

The better hypothesis will have a lower value of χ^2 and a greater value P, where P denotes the probability of obtaining the 'observed' deviations from uniform distribution or larger ones. The probability P depends upon the number of cells chosen. Therefore, to compare two competing hypotheses, the same division must be used.

The application of the χ^2 test to the 'observations' $W(x_m)$ eliminates an arbitrary action which is a serious and well-known drawback of the χ^2 test, when applied to the original observations x_m. The expected contents of the classes depend upon the distribution. Therefore certain classes, as a rule the first and the last, must be chosen such that the expected number is not too small, otherwise χ^2 becomes very large. In our case, no cell differs from any other and no arbitrary combination of cells is needed. We can choose $k = n$. The mean number of points in each cell will then be one and

$$\chi_0^2 = \sum_{\nu=1}^{n}(a_\nu - 1)^2, \quad \chi_1^2 = \sum_{\nu=1}^{n}(b_\nu - 1)^2. \tag{18'}$$

This choice removes another drawback of the χ^2 method: different classifications used for the same observations lead to different shapes of the distribution and therefore to different values of χ^2. Here the classification is prescribed once and for all.

Another comparison between theory and 'observation' may be based on the fact that different sets of points a_ν and b_ν have different probabilities. The probability that a_ν points will fall within the cell ν $(= 1, 2, \ldots, k)$ is

$$\frac{n!}{a_1!\,a_2!\ldots a_k!}\left(\frac{1}{k}\right)^{a_1+a_2+\ldots+a_k},$$

where

$$a_1+a_2+\ldots+a_k = n.$$

Since the factor $n!\,k^{-n}$ is constant, it is sufficient to investigate

$$\Pi = \frac{1}{\prod_{\nu=1}^{k} a_\nu!}. \tag{19}$$

Of course $\Pi < P$, as the latter probability applies to the 'observed' deviation or larger ones. The statement 'The probability for a_ν points to be contained in the cell ν is proportional to Π' may be inverted according to Bayes's principle. Therefore, Π is proportional to the probability that the distribution of points is rectangular, i.e. that h_0 is a good hypothesis.

The question for which set $\tilde{a}_\nu$ the probability Π is maximum, is the starting-point of the classical relation between entropy and probability. For large n the most probable set of points is the one which has the same number of points in each cell, i.e.

$$\tilde{a}_1 = \tilde{a}_2 = \ldots = \tilde{a}_k = \frac{n}{k}. \tag{20}$$

Let us call $\Pi_{\max.}$ the probability which corresponds to this distribution. The probability of the hypothesis h_0 will be greater, equal to, or less than the probability of h_1, if

$$\frac{\Pi_0}{\Pi_{\max.}} \gtreqless \frac{\Pi_1}{\Pi_{\max.}}. \tag{21}$$

The relative probability of both hypotheses will be Π_1/Π_0 or Π_0/Π_1, depending on whether

$$\Pi_0 \gtrless \Pi_1.$$

As these probabilities depend on the number of cells, the same division must be used to compare two competing hypotheses. We can choose $k = n$. Then $\Pi_{\max.} = 1$ and we can use Π_0 as test.

The entropy test (21) is closely related to the χ^2 test (18). This classical relation can be obtained in the following simple way: if q is the constant probability of a point falling within a given cell, then for n observations the expected

number of points within a cell is nq. But the 'observed' numbers a_ν will differ from the expected number by ϵ_ν so that

$$a_\nu = nq + \epsilon_\nu,$$

where

$$\sum_{\nu=1}^{k} \epsilon_\nu = 0.$$

The quotient (21) becomes

$$\frac{\Pi_0}{\Pi_{\max.}} = \prod_{\nu=1}^{k} \frac{(nq)!}{(nq+\epsilon_\nu)!}.$$

When n is large, each factor becomes, by application of Stirling's formula,

$$\frac{nq!}{(nq+\epsilon_\nu)!} = \left(\frac{e}{nq}\right)^{\epsilon_\nu} \exp\left[-(nq+\epsilon_\nu+\tfrac{1}{2})\ln\left(1+\frac{\epsilon_\nu}{nq}\right)\right].$$

Expansion of the logarithm leads to

$$\frac{nq!}{(nq+\epsilon_\nu)!} = \left(\frac{e}{nq}\right)^{\epsilon_\nu} \exp\left[-\epsilon_\nu - \frac{\epsilon_\nu}{2nq} - \frac{\epsilon_\nu^2}{2nq}\right].$$

According to the meaning of ϵ_ν we obtain

$$\frac{\Pi_0}{\Pi_{\max.}} = \exp\left[-\frac{1}{2}\sum_{\nu=1}^{k}\left(\frac{a_\nu - nq}{nq}\right)^2\right],$$

whence

$$\frac{\Pi_0}{\Pi_{\max.}} = e^{-\frac{1}{2}\chi^2}. \tag{22}$$

Therefore, when n is large, the entropy test becomes identical with the χ^2 test. This result was derived by Neyman & Pearson (1928), when they showed that the χ^2 test followed from their 'likelihood ratio' method of approach.

Neither test will give an answer, if the number of points a_ν assigned by h_0 to the cell ν is equal to the number of points b_λ assigned by h_1 to the cell λ, $a_\nu = b_\lambda$, where for any ν $(= 1, 2, \ldots, k)$ it is possible to find a λ $(= 1, 2, \ldots, k)$, such that not all $\lambda = \nu$. An example of this occurring is shown in Table 1, col. C and F. The reason for the failure is that we do not make use of the actual position of the observed points within the cells. We only ask in which cell they are situated. Although in such a case the tests do not show any difference between the arrangements a_ν and b_λ, some conclusions might be drawn from such 'observations'. If the number of points falling in the first few cells and also in the last few cells is disproportionately large, and if there is a deficiency in the middle cells (Table 1, col. D), we have to conclude that the distribution h_0 is too concentrated or that we have chosen too small a value for the constant which depends only on the standard deviation. If the number of points in the cells at either end is small (Table 1, col. E), the inverse inferences follow. These considerations may give a hint about the choice of an alternative hypothesis.

To illustrate the above methods, let us take the fictitious example given by Pearson (1938) in his Fig. 2, p. 136. He arranges $n = 10$ points in $k = 10$ cells and considers six sets A, B, ..., F, given in Table 1.

Let us suppose that these six sets are the results of six different hypotheses applied to the same observations. The χ^2 test leads to

$$P_A > P_D > P_C = P_F > P_B = P_E.$$

The probabilities of the various columns give the same ordering

$$\Pi_{\max.} = \Pi_A > \Pi_D > \Pi_C = \Pi_F > \Pi_B = \Pi_E.$$

The most probable set contains one point in each cell (set A). It is not possible to decide whether C is more probable than F, and whether B is more probable than E.

Table 1. *Pearson's set*

Class	A	B	C	D	E	F
0·0–0·1	1	2	0	2	0	0
0·1–0·2	1	3	0	1	0	1
0·2–0·3	1	2	0	2	1	2
0·3–0·4	1	1	0	0	2	2
0·4–0·5	1	0	1	0	2	0
0·5–0·6	1	1	2	1	3	1
0·6–0·7	1	0	1	0	1	0
0·7–0·8	1	1	2	1	1	0
0·8–0·9	1	0	2	1	0	2
0·9–1·0	1	0	2	2	0	2
χ^2	0	10	8	6	10	8
P	1	0·350	0·534	0·740	0·350	0·534
Π	1	$\frac{1}{24}$	$\frac{1}{18}$	$\frac{1}{8}$	$\frac{1}{24}$	$\frac{1}{18}$

The χ^2 and the entropy test are based upon the same data. But the results reached are incomplete, as artificialities are introduced by the classification of the 'observations' into the arbitrary cells. The actual position of the points within the cells is not used. The set A shows that these tests may be misleading in still another way. Each cell in set A contains exactly the expected number. But it would be false to conclude that the hypothesis is true, since the actual positions of the points within the cells might differ from the ideal positions.

Let us suppose we know these positions. It might then happen that the difference between the observed and the ideal positions of the points is smaller for a set K than for a set L, even if the differences between the actual and the theoretical number of points is larger for K than for L.

It is now our task to assign a meaning to the term ideal position and to define a measure of the differences between the 'observed' and the ideal set.

3. The mth point test

The ideal position of n points, distributed with uniform probability over the interval 0, 1, is such that the distances between consecutive pairs of points are equal. But there are a number of ways of distributing n points equidistantly over the interval 0, 1. E. S. Pearson, in the preceding note, suggests that

$$\hat{y}_m = \frac{2m-1}{2n} \tag{23}$$

might be used as the ideal position of the mth point. However, as y is a statistical variable, we should represent it by an average, to choose which we must consider the distribution of the mth point.* Any observation chosen at random has the same probability of falling on any position y within the interval. But for the mth point this probability depends upon y and m. The initial distribution w of the variable $W(x) = y$ is constant. According to (8)

$$w(y) = 1 \quad (0 \leqq y \leqq 1).$$

The probability of obtaining a point equal to or less than y is y. According to (2) the distribution of the mth point is

$$\mathfrak{w}_m(y, n) = \binom{n}{m} m y^{m-1}(1-y)^{n-m}. \tag{24}$$

The distribution (24) is of Karl Pearson's Type I. For $m = 1$ (and $m = n$) the distribution will only decrease (increase). The distribution of the mth point is equal to the distribution of the $(n-m+1)$th point. If we replace y by $1-y$

$$\mathfrak{w}_{n-m+1}(1-y, n) = \mathfrak{w}_m(y, n). \tag{24'}$$

The most probable position $\tilde{y}$ of the mth point is given by

$$\frac{n-m}{1-y} = \frac{m-1}{y},$$

which leads to

$$\tilde{y}_m = \frac{m-1}{n-1}. \tag{25}$$

For given values of m the median position can be obtained from the tables of the Incomplete Gamma Function. To find the arithmetic mean $\bar{y}$ and the control curves, it is necessary to have the moments M_k of (24). They are

$$M_k = \binom{n}{m} m \int_0^1 y^{m+k-1}(1-y)^{n-m}\, dy.$$

According to the well-known properties of the Gamma function

$$M_k = \frac{n!}{(m-1)!\,(n-m)!} \frac{(m+k-1)!\,(n-m)!}{(n+k)!} = \frac{n!}{(n+k)!} \frac{(m+k-1)!}{(m-1)!}.$$

* [The distribution of a ranked individual sampled from a rectangular population, and the moments of this distribution, were obtained by Karl Pearson in the first of two papers (1931, p. 390, and 1932) dealing with ranked variates. Ed.]

Therefore $$M_{k+1} = M_k \frac{m+k}{n+k+1}. \tag{26}$$

For $k = 1$ the arithmetic mean of the mth point is

$$\bar{y}_m = \frac{m}{n+1}, \tag{27}$$

and for $k = 2$ the second moment is

$$M_2 = \bar{y}_m \frac{m+1}{n+2}.$$

Finally, the variance $$\sigma_m^2 = \frac{m(n-m+1)}{(n+1)^2(n+2)}. \tag{28}$$

These formulae apply also to the cases $m = 1$ and $m = n$. The standard deviation of the mth value may be written

$$\sigma_m = \sqrt{\frac{\bar{y}_m(1-\bar{y}_m)}{n+2}}. \tag{28'}$$

This formula differs slightly from the general expression (3), and leads to an unexpected result: as we approach the centre from either side, the precision of the mth point decreases. The precision of the mth point will be a minimum for $m = \frac{1}{2}n+1$, if n is even, and for $m = \frac{1}{2}(n+1)$ if n is odd. The values of $\sigma_m\sqrt{(n+2)}$ are given in Table 2.

Table 2. *Standard deviation of the mth point*

$\bar{y}_m$	$\bar{y}_m$	$\sigma_m\sqrt{(n+2)}$
0·05	0·95	0·21794
0·10	0·90	0·30000
0·15	0·85	0·35707
0·20	0·80	0·40000
0·25	0·75	0·43301
0·30	0·70	0·45826
0·35	0·65	0·47697
0·40	0·60	0·49000
0·45	0·55	0·49750
0·50	0·50	0·50000

We must now decide whether to use the mean (27) or the mode (25) as the ideal position of the mth point. The modes of the first and of the last points are 0 and 1 respectively, whereas the corresponding means are

$$\bar{y}_1 = \frac{1}{n+1}, \quad \bar{y}_n = 1 - \frac{1}{n+1}.$$

As the 'observations' $W_0(x_1)$ and $W_0(x_n)$ of the first and the last point differ from 0 and 1, the arithmetic mean is to be preferred.

Formula (27) gives the ideal position and therefore the theoretical numbers of points in each cell which can be compared with the 'observed' numbers. This method leads to an improvement of the tests (18) and (21) where the choice of the cells was still arbitrary and where the actual position of each point was not taken into account.

Besides comparing the uniform distribution with the 'observed' position of the points we can use the corresponding cumulative frequency. The probability scale y is plotted as abscissa and m as ordinate. We count the number of points below m. The mean position $\bar{y}$ of the mth point becomes a straight line differing from the diagonal which represents the modal positions. The figure opposite traces, for $n = 20$, the mean, the modal and Pearson's position of the mth point given by (23).

In the same way we plot the 'observed' points obtained by h_0, h_1, ... These probability points $W_0(x_m)$, $W_1(x_m)$ will be scattered about the straight line. Usually the area between the observed cumulative frequency curve, the ordinate and the parallel to the abscissa is kept equal to the corresponding area for the theoretical curve. Since for the present problem no constants have to be determined, we have no way of enforcing this equality. The area J bounded by the diagonal straight line through the points with the co-ordinates $m/(n+1)$, m $(m = 1, 2, \ldots, n)$, the length $1/(n+1)$ to $n/(n+1)$ of the abscissa axis and the two parallels to the ordinate axis, is

$$J = \tfrac{1}{2}(n-1). \tag{29}$$

This might differ from the area $J^{(0)}$ of the $n-1$ 'observed' trapezes

$$y_m,\ m,\ m+1,\ y_{m+1} \quad (m = 1, 2, \ldots, n-1).$$

As y_m are the 'observed' points

$$\begin{aligned} J^{(0)} &= \sum_{m=1}^{n-1} (m+\tfrac{1}{2})\,(y_{m+1}-y_m) \\ &= \sum_{2}^{n} y_m(m-1) - \sum_{1}^{n-1} m y_m + \frac{1}{2}\sum_{2}^{n} y_m - \frac{1}{2}\sum_{1}^{n-1} y_m \\ &= n y_n - \sum_{1}^{n} y_m + \tfrac{1}{2}(y_n - y_1). \end{aligned}$$

If we replace each value y_m by its expectation from (27), we get

$$\begin{aligned} \bar{J}^{(0)} &= \frac{1}{n+1}\{n^2 - \tfrac{1}{2}[n(n+1)] + \tfrac{1}{2}(n-1)\} = \frac{n^2-1}{2(n+1)} \\ &= \tfrac{1}{2}(n-1), \end{aligned} \tag{29'}$$

as it ought to be. The 'observed' area is not equal to the theoretical area, but its expectation is. It might happen that the numerical value of $J^{(0)}$ is very close to J as a result of compensating deviations. Therefore this numerical comparison can be used as a test only in connexion with the graph of the 'observed' and theoretical cumulative histogram of the mth points.

To control the agreement between the 'observed' cumulative histogram and the ideal straight line we use two control curves through the points $\bar{y} \mp \sigma$, m, where $\bar{y}$ is given by (27) and σ by (28'). They are traced in the figure for $n = 20$.

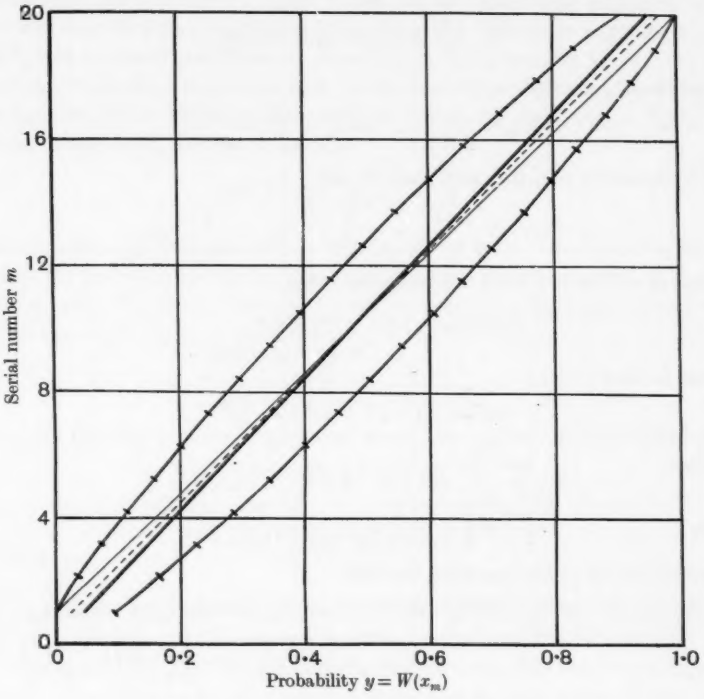


Fig. 1. Control curve for uniform distribution.

——— Mean mth point $\bar{y}$ (27) ——— Modal mth point $\tilde{y}$ (25)
– – – – Pearson's mth point $\hat{y}$ (23) |—|—|—| Control curves

It is interesting to calculate the area A bounded by these two curves and to compare it with the area J. If we consider m as abscissa and $\bar{y}+\sigma$ and $\bar{y}-\sigma$ as ordinates we have, for n sufficiently large,

$$A = \int_1^n [\bar{y}+\sigma-(\bar{y}-\sigma)]\,dm$$
$$= 2\int_1^n \sigma\,dm.$$

If we introduce $\bar{y}$ as variable of integration we obtain from (28')

$$A = \frac{2(n+1)}{\sqrt{(n+2)}} \int_{1/(n+1)}^{n/(n+1)} \sqrt{\{\bar{y}(1-\bar{y})\}}\,d\bar{y}.$$

The transformation

$$\bar{y} = \sin^2 t, \quad 1-\bar{y} = \cos t, \quad d\bar{y} = 2\sin t \cos t\, dt,$$

leads, as is well known, to

$$\frac{\sqrt{(n+2)}}{n+1} A = \tfrac{1}{4}[t - \sin t \cos t]_{2t_0}^{2t_1}.$$

The limits are given by

$$\sin t_0 = \sqrt{\frac{1}{n+1}}, \quad \sin t_1 = \sqrt{\left(1-\frac{1}{n+1}\right)}.$$

For the expansion of t_0 it is sufficient to put

$$t_0 = \arcsin\sqrt{\frac{1}{n+1}} = \frac{1}{\sqrt{(n+1)}} - \frac{1}{6(n+1)\sqrt{(n+1)}},$$

provided $(n+1)^2 \gg 1$. Under the same condition,

$$\begin{aligned}\arcsin\sqrt{(1-x^2)} &= \arccos x \\ &= \tfrac{1}{2}\pi - \arcsin x,\end{aligned}$$

becomes for any $|x| < 1$

$$\arcsin\sqrt{(1-x^2)} = \tfrac{1}{2}\pi - x + \tfrac{1}{6}x^3.$$

Therefore

$$t_1 = \frac{\pi}{2} - \frac{1}{\sqrt{(n+1)}} + \frac{1}{6\sqrt{(n+1)}\,(n+1)}$$

so that

$$\frac{t_1 - t_0}{2} = \frac{\pi}{4} - \frac{1}{\sqrt{(n+1)}} + \frac{1}{6(n+1)\sqrt{(n+1)}}.$$

The second factor in the brackets becomes

$$\begin{aligned}2\sin t_0\sqrt{(1-\sin^2 t_0)}\,(\cos^2 t_0 - \sin^2 t_0) &= 2\sin t_0\sqrt{(1-\sin^2 t_0)}\,(1-2\sin^2 t_0) \\ &= 2\sqrt{\left\{\frac{1}{n+1}\left(1-\frac{1}{n+1}\right)\right\}}\left(1-\frac{2}{n+1}\right) \\ &= \frac{2\sqrt{n}\,(n-1)}{(n+1)^2}.\end{aligned}$$

In the same way

$$\sin 2t_1\sqrt{(1-\sin^2 2t_1)} = 2\sqrt{\left\{\frac{n}{n+1}\left(1-\frac{n}{n+1}\right)\right\}}\left(1-\frac{2n}{n+1}\right) = -\frac{2\sqrt{n}\,(n-1)}{(n+1)^2}.$$

Finally, we reach the area bounded by the control curves

$$A = \frac{n+1}{\sqrt{(n+2)}}\left(\frac{\pi}{4} + \frac{(n-1)\sqrt{n}}{(n+1)^2} - \frac{1}{\sqrt{(n+1)}} + \frac{1}{6(n+1)\sqrt{(n+1)}}\right).$$

According to (29) the ratio of the area bounded by the control curves to the area of the cumulative histogram

$$\frac{A}{J} = \frac{n+1}{n-1}\frac{2}{\sqrt{(n+2)}}\left(\frac{\pi}{4} + \frac{(n-1)\sqrt{n}}{(n+1)^2} - \frac{1}{\sqrt{(n+1)}} + \frac{1}{6(n+1)\sqrt{(n+1)}} \cdots\right), \qquad (30)$$

converges towards zero as $1:\sqrt{n}$.

The properties (29′) and (30) of the cumulative histogram of the positions of the mth points allow for the comparison of the 'observations' $W_0(x_m)$ and the ideal points $m/(n+1)$. It will often be sufficient to inspect the deviations between the 'observations' and theory to judge which set of 'observed' points is closer, on the whole, to the theoretical positions. It seems legitimate to prefer a hypothesis h_0 if the control area contains more points for h_0 than for h_1.

In order to secure a numerical test, we can introduce the mean $\mathfrak{S}^2$ of the sum of the squares of the differences between the 'observed' values $y_m = W_0(x_m)$ and the mean positions $\bar{y}_m = m/(n+1)$. Take

$$\mathfrak{S}^2 = \frac{k}{n} \sum_{m=1}^{n} \left(y_m - \frac{m}{n+1} \right)^2, \tag{31}$$

where the value of the constant k will be specified later. One extreme for (31) would be to have the theory hold for every point. Then the value of the sum would be zero. The other extremes would be when all points are concentrated either at the origin, zero, or at the end, 1. In the first case

$$\frac{1}{n(n+1)^2} \sum_{m=1}^{n} m^2 = \frac{2n+1}{6(n+1)}.$$

The second case leads to the same value, since $\Sigma m = \frac{1}{2} n(n+1)$ and therefore

$$\frac{1}{n} \sum_{m=1}^{n} \left(1 - \frac{m}{n+1} \right)^2 = 1 - 1 + \frac{1}{n(n+1)^2} \sum_{m=1}^{n} m^2.$$

Therefore

$$0 < \mathfrak{S}^2 < k \frac{2n+1}{6(n+1)} < \frac{k}{3}. \tag{32}$$

In order to draw conclusions from an observed value $\mathfrak{S}^2$ we have to calculate its expectation $\overline{\mathfrak{S}^2}$. We will determine k in such a way that $\overline{\mathfrak{S}^2}$ is independent of n.

A test similar to (31), but serving another purpose, has been introduced for the usual distributions by H. Cramér (1928) and R. von Mises (1931). When applied to uniform distributions, this ω^2 test leads to the use of $\hat{y}_m$ of (23) instead of the mean value $\bar{y}_m$. For this test the sum of the deviations is zero which does not hold in our case. The expectation $\overline{\mathfrak{S}^2}$ of $\mathfrak{S}^2$ is

$$\overline{\mathfrak{S}^2} = \frac{k}{n} \left(\sum_{m=1}^{n} \left[\overline{y_m^2} - \frac{2m\bar{y}_m}{n+1} + \frac{m^2}{(n+1)^2} \right] \right).$$

The first two sums are obtained from (27) and (26). Therefore

$$\overline{\mathfrak{S}^2} = \frac{k}{n} \left(\frac{\sum_1^n m(m+1)}{(n+1)(n+2)} - \frac{\sum_1^n m^2}{(n+1)^2} \right)$$

$$= \frac{k}{n(n+1)} \left(\frac{\sum_1^n m}{n+2} - \frac{\sum_1^n m^2}{(n+1)(n+2)} \right).$$

The introduction of the sum of the powers of the natural numbers leads to

$$\overline{\mathfrak{S}^2} = \frac{k}{2(n+2)}\left(1-\frac{2n+1}{3n+3}\right) = \frac{k}{6(n+1)}.$$

Taking $k = 6(n+1)$ we propose therefore, as test of a hypothesis h_0, the coefficient

$$\mathfrak{S}^2 = \frac{6(n+1)}{n}\sum_{m=1}^{n}\left(y_m - \frac{m}{n+1}\right)^2, \tag{33}$$

which, according to (32), can assume values between zero and $2n+2$, and has for expectation the value

$$\overline{\mathfrak{S}^2} = 1. \tag{34}$$

Of two competing hypotheses, the one with the smaller value of $\mathfrak{S}^2$ is to be preferred.

The $\mathfrak{S}^2$ criterion does not introduce any arbitrary classification. It makes use of all observations. Besides the probabilities $W(x_m)$ corresponding to the observed values x_m no new calculations are needed. The test has a clear meaning and its application is simple. This is due to the fact that it is a natural consequence of the probability integral transformation.

Summary

We propose the following procedure for testing statistical hypotheses: The constants for competing hypotheses, having the same number of constants, are determined in such a way that their precisions are approximately the same. Then we calculate the probabilities $W_0(x_m)$, $W_1(x_m)$, ..., and their respective control curves. We trace $W(x_m)$, $x_m \mp \sigma_m$ and compare it with the observed frequency curve. If neither the classical tests nor the control curves indicate a clear superiority of one of the hypotheses we consider the probabilities as 'observations' and plot the corresponding points on the y axis. We now compare, by formulae (9)—(17), the 'observations' with the theoretical uniform distribution and apply the χ^2 and entropy test of formulae (18) and (21), respectively. If necessary, we repeat these tests in such a way that the actual position of each point is taken into account. Formula (18′) gives a value of χ^2 which is independent of the classification. Then we plot the cumulative frequency of the 'observations', which is compared with the straight line (27) and controlled by the values given in Table 2. The final test is given by (33).

REFERENCES

CRAMÉR, H. (1928). On the composition of elementary errors. *Skand. Aktuar. Tidskr.* **11**, 13–17, 141–80.

GUMBEL, E. J. (1935). Les valeurs extrêmes des distributions statistiques. *Ann. Inst. Henri Poincaré*, **4**, fasc. 2, 115–58.

—— (1936). La précision de la moyenne arithmétique et de la médiane. *Aktuar. Vedy*, **6**, no. 4, 145–54.

—— (1937). Les intervalles extrêmes entre les émissions radioactives. I. *J. Phys. Radium*, série VII, **8**, no. 8, pp. 321–29; II, **446–52**.

—— (1939). Les valeurs de position d'une variable aléatoire. *C.R. Acad. Sci., Paris*, **208**, 147–9.

MISES, R. VON (1931). *Wahrscheinlichkeitsrechnung*, 316–35. Leipzig.

NEYMAN, J. (1937). 'Smooth test' for goodness of fit. *Skand. Aktuar. Tidskr.* 149–99.

NEYMAN, J. & PEARSON, E. S. (1928). On the use and interpretation of certain test criteria for purposes of statistical inference. *Biometrika*, **20A**, 263–94.

PEARSON, E. S. (1938). The probability integral transformation for testing goodness of fit and combining independent tests of significance. *Biometrika*, **30**, 134–48.

PEARSON, KARL (1933). On a method of determining whether a sample of size n supposed to have been drawn from a parent population having a known probability integral has probably been drawn at random. *Biometrika*, **25**, 379–410.

—— (1934). On a new method of determining goodness of fit. *Biometrika*, **26**, 425–42.

PEARSON, KARL & PEARSON, M. V. (1931, 1932). On the mean character and variance of a ranked individual, and on the mean and variance of the intervals between ranked individuals. Part I: *Biometrika*, **23**, 364–97. Part II: *Biometrika*, **24**, 203–79.

THE RANGE IN RANDOM SAMPLES

By H. O. HARTLEY

1. Introduction

If the observations x_ν $(\nu = 1, 2, \ldots, n)$ of a random sample are arranged in ascending order of magnitude $(x_{\nu+1} > x_\nu)$ the range w in such samples is defined as the distance between the two extreme observations

$$w = x_n - x_1.$$

It may therefore be regarded as a measure of the variability or dispersion among the observations of the sample. Theoretically its efficiency in the sense defined by R. A. Fisher is, as a rule, much inferior to that of the standard deviation. Moreover, extensive investigations have shown that its random sampling distribution is markedly dependent on the parental population (E. S. Pearson, 1926). For large samples x_ν drawn from a parental distribution $f(x)$ the extreme values x_1 and x_n will lie right inside the lower and upper tail of $f(x)$, and in practice it is only in exceptional cases that the exact shape of $f(x)$ has been established to such a degree of accuracy that the resulting distribution of w can be trusted for large n. In most cases the use of the range must therefore be limited to small samples, say with $2 \leqslant n \leqslant 20$.

Large numbers of small samples may often be used with advantage when the mean range is calculated as an estimate of the standard deviation of the population (Pearson & Haines, 1935). Although theoretically such an estimate is not efficient and certainly not sufficient, it is nevertheless of considerable importance in many fields of application because of its simplicity. Statistical control charts in industrial quality control make extensive use of it, and more recently the range has been applied to investigations in gunnery.

In some fields of application a disadvantage may arise from the fact that the range is an inexact statistic; its random sampling distribution depends on the standard deviation of the parent. This applies in particular to the analysis of small samples in biological experiments. The tendency of modern small sample theory has been to replace such statistics by what are called exact statistics, obtained by substituting for the unknown standard deviation of the parent an estimate calculated from an independent sample. This particular process of reaching exact statistics has sometimes been referred to as 'Studentization'. A general theory of this process will be given in a further paper which it is hoped to publish in this journal, where it will be shown how estimates of scale parameters in general, and of the range in particular, may be converted into exact statistics. In this paper, however, we deal with the case where the standard deviation of the parent is known. Indeed, it is this dependence of the random sampling distribution of range on the scale parameter of the parent which makes it possible to estimate its efficiency as an estimate theoreof.

The question of grouping has been a subject of investigation in the case of the sample standard deviation; we shall here deal with the effect of grouping on the range, a problem which has so far received, we believe, no attention whatsoever.* As practical examples of the occurrence of grouping we may quote three instances:

(*a*) The rounding off of data for convenience of recording and analysis.

(*b*) The recording of data to the nearest unit of measurement. Where the technique is of low accuracy (see e.g. Tildesley, 1940) the unit of measurement will be comparable in magnitude with the standard deviation of the actual data.

(*c*) The analysis of data which are classified in categories. In such cases we may often find that the original data are unobtainable so that group frequencies are the only material available for an analysis.

It will also be shown how the random sampling distribution of the range in grouped samples provides a suitable approach to that of the true range (ungrouped range in sample) on which extensive work has been done in earlier papers published in *Biometrika*. The mathematical formulae developed in this paper make this complex distribution amenable to a tabulation. For the case of normal samples, the work has actually been carried out and the resulting tables of the probability integral are given and discussed elsewhere in the present issue of this journal.

2. THE DISTRIBUTION OF THE RANGE IN A GROUPED SAMPLE

Let us denote by $x_1, \ldots, x_n$ the observations in a random sample drawn from the parental distribution $f(x)$† and arranged in ascending order of magnitude. This sample is now classified in groups or categories of constant length h with equidistant end-points

$$\ldots, \xi - kh, \ldots, \xi - h, \xi, \xi + h, \ldots, \xi + kh, \ldots, \tag{1}$$

covering the whole x scale from $-\infty$ to $+\infty$. Let us denote by ξ_n and ξ_1 the respective *centres* of the categories containing x_n and x_1. Then the problem is to find the random sampling distribution of the range in a grouped sample, i.e. of $\xi_n - \xi_1$. The mean of this distribution is of particular interest. Obviously this statistic can only assume values which are multiples of the group interval h and is therefore discontinuous. Like the distribution of the 'ungrouped' or true range it depends on the standard deviation σ of the parent $f(x)$. In addition, it depends

* The effect of grouping seems to be of some importance in researches on the technique of anthropological measurement (Tildesley, 1940), where some of the results given below have already been applied before this paper had gone to press.

† We shall deal here with a parental population represented by a 'piecewise continuous' distribution function $f(x)$. A function is called 'piecewise continuous' for $-\infty < x < +\infty$ if in any closed interval of x the function $f(x)$ is continuous apart from a finite number of ordinary discontinuities. If the actual range of the variate is bounded we simply define $f(x) \equiv 0$ outside this range. Moreover, we assume that $f(x)$ has contact of at least second order at $\pm\infty$. It is easy to see how our results may be generalized to cover distribution functions with singularities.

on the category width h and on the position of the category midpoints relative to the population mean X. Of these parameters only h will in practice be known. Methods to eliminate σ are to be given in a separate paper whilst the elimination of X is dealt with in the section on randomized grouping (6).

It will be convenient to use the following notation:

$$\int_i^j = \int_{\xi+ih}^{\xi+jh} f(x)\,dx, \quad \int_{-\infty}^{i} = \int_{-\infty}^{\xi+ih} f(x)\,dx, \quad \int_i^{\infty} = \int_{\xi+ih}^{\infty} f(x)\,dx.$$

Let us now find the chance that $\xi_n - \xi_1$ is at most $(m-1)h$, and that in addition

$$\xi_1 = \xi + (i+\tfrac{1}{2})h$$

for a particular value of i. This chance is given by

$$\left(\int_i^{i+m}\right)^n - \left(\int_{i+1}^{i+m}\right)^n. \tag{2}$$

The first term in (2) represents the probability for all x_ν to lie between $\xi+ih$ and $\xi+(i+m)h$. From this we have to deduct the chance for all x_ν to lie between $\xi+(i+1)h$ and $\xi+(i+m)h$ which is given by the second term of (2). In taking the difference we are therefore left with the chance for the occurrence of a sample completely contained in the interval $\xi+ih$ to $\xi+(i+m)h$ but with at least one of the x_i lying between $\xi+ih$ and $\xi+(i+1)h$. This proves that (2) represents the required chance. Now, since all samples may be classified with regard to their lowest category, the probability for $\xi_n-\xi_1$ to be at most $(m-1)h$ is given by summation over all i of the expression (2). If we denote this probability by $P(n,h,m-1,\xi)$ we find

$$P(n,h,m-1,\xi) = \sum_{i=-\infty}^{+\infty} \left\{\left(\int_i^{i+m}\right)^n - \left(\int_{i+1}^{i+m}\right)^n\right\}. \tag{3}$$

With equation (3) we have reached a formal representation of the random sampling distribution of $\xi_n-\xi_1$. Its evaluation is a simple matter for large group intervals h and for parental distributions $f(x)$ with a tabulated probability integral $\int f(x)\,dx$. If we were to take the trouble of tabulating the probability integral (3) we should obtain the mean of $\xi_n-\xi_1$ as a by-product from a summation of (3). It will be shown in the next section that this summation, if carried out analytically, produces a very simple formula for this mean.

3. The mean range in a grouped sample

To find the mean of the distribution it is convenient to extend the summation in (3) from some finite negative value $i=-j$ up to $+\infty$. By choosing j sufficiently large the resulting error may be made negligible. We introduce

$$p_{-j}(m-1) = \sum_{i=-j}^{\infty} \left[\left(\int_i^{i+m}\right)^n - \left(\int_{i+1}^{i+m}\right)^n\right], \tag{4}$$

and find for the difference between $P(m-1)$ and $p_{-j}(m-1)$*

$$|P(m-1)-p_{-j}(m-1)| \leqslant \sum_{i=-\infty}^{-j} n\int_{i}^{i+1} = n\int_{-\infty}^{-j+1}, \tag{5}$$

for all m and j. To find the mean of $\xi_n - \xi_1$ we must first note the probability for this statistic to be exactly equal to mh, where $m = 0, 1, 2, \ldots$. Denoting this probability by $\phi(m)$ we have from the definition of $P(m)$

$$\phi(m) = P(m) - P(m-1).$$

If we denote the mean of $\xi_n - \xi_1$ by Ξ we have by definition

$$\Xi = h\sum_{k=0}^{\infty}\phi(k)\,k$$

$$= h\lim_{m\to\infty}\{(m+1)\,P(m) - S_m\}, \tag{6}$$

where $$S_m = (m+1)\phi(0) + m\phi(1) + \ldots + \phi(m)$$

or $$S_m = P(0) + P(1) + \ldots + P(m). \tag{7}$$

To find Ξ let us first consider the second term in formula (6). We have from equations (7) and (5)

$$S_m = p_{-j}(0) + \ldots + p_{-j}(m) + \epsilon_1, \tag{8}$$

where $$|\epsilon_1| \leqslant n(m+1)\int_{-\infty}^{-j+1}, \tag{9}$$

for all j and m. Now from the definition of $p_{-j}(k)$ we find

$$\sum_{k=0}^{m} p_{-j}(k) = \sum_{k=0}^{m}\left\{\sum_{i=-j}^{\infty}\left[\left(\int_{i}^{i+k+1}\right)^n - \left(\int_{i+1}^{i+k+1}\right)^n\right]\right\}$$

$$= \sum_{k=0}^{m}\left\{\left(\int_{-j}^{-j+k+1}\right)^n + \sum_{i=-j}^{\infty}\left[\left(\int_{i+1}^{i+k+2}\right)^n - \left(\int_{i+1}^{i+k+1}\right)^n\right]\right\}$$

$$= \sum_{k=0}^{m}\left(\int_{-j}^{-j+k+1}\right)^n + \sum_{i=-j}^{\infty}\left(\int_{i+1}^{i+m+2}\right)^n.$$

Putting now $m = 2j$, we have

$$\sum_{k=0}^{2j} p_{-j}(k) = \sum_{i=-j+1}^{j+1}\left(\int_{-\infty}^{i}\right)^n + \sum_{i=-j+1}^{j+1}\left(\int_{i}^{\infty}\right)^n + \epsilon_2 + \epsilon_3, \tag{10}$$

where it is easily seen that

$$|\epsilon_2| \leqslant 2jn\int_{-\infty}^{-j}, \quad |\epsilon_3| \leqslant 2jn\int_{j+2}^{\infty} + \sum_{i=j+1}^{\infty}(i-j)\int_{i}^{i+1}. \tag{11}$$

Finally, we want to replace in formula (6) the first term

$$(2j+1)\,P(2j) \quad \text{by} \quad (2j+1). \tag{12}$$

* In this section we deal with fixed group intervals and a fixed sample size n so that we drop the arguments n, h and ξ.

The resulting error is easily estimated. We have from (5)

$$(2j+1)(1-P(2j)) = (2j+1)(1-p_{-j}(2j))+\epsilon_4$$

where
$$|\epsilon_4| \leqslant (2j+1)\,n\int_{-\infty}^{-j+1}. \tag{13}$$

Moreover, according to the definition (4) we may write

$$\begin{aligned}(2j+1)\,p_{-j}(2j) &= (2j+1)\left\{\sum_{i=-j}^{\infty}\left(\int_i^{i+2j+1}\right)^n-\left(\int_{i+1}^{i+2j+2}\right)^n\right\}\\ &\quad+(2j+1)\left\{\sum_{i=-j}^{\infty}\left(\int_{i+1}^{i+2j+2}\right)^n-\left(\int_{i+1}^{i+2j+1}\right)^n\right\}\\ &= (2j+1)\left(\int_{-j}^{+j+1}\right)^n+\epsilon_5,\end{aligned}$$

where
$$|\epsilon_5| \leqslant (2j+1)\sum_{i=-j}^{\infty} n\int_{i+2j+1}^{i+2j+2} \leqslant n(2j+1)\int_{j+1}^{\infty}, \tag{14}$$

so that finally we have

$$(2j+1)(1-p_{-j}(2j)) = -\epsilon_5+\epsilon_6,$$

with
$$\epsilon_6 \leqslant (2j+1)\,n\left(\int_{j+1}^{\infty}+\int_{-\infty}^{-j}\right). \tag{15}$$

The error terms ϵ_1, ϵ_2, ϵ_3, ϵ_4, ϵ_5 and ϵ_6 are of the form

$$cj\int_j^{\infty}; \quad cj\int_{-\infty}^{-j} \quad \text{or} \quad \sum_{i=j}^{\infty}(i-c)\int_i^{i+1}.$$

It is easy to see that the above terms tend to 0 as $j\to\infty$. To prove this for the first term we write

$$j\int_j^{\infty} \leqslant \left(\frac{\xi+jh}{h}+c\right)\int_j^{\infty} < c\int_j^{\infty}+\frac{1}{h}\int_{\xi+jh}^{\infty} xf(x)\,dx,$$

which tends to 0 as $f(x)$ has contact of order 1 at $+\infty$. The proof for the other terms is identical. For sufficiently large j we can therefore use the approximations given by (10) and (12) and transform equation (6) into the convenient form

$$\Xi = h\lim_{j\to\infty}\left\{(2j+1)-\sum_{i=-j+1}^{j+1}\left(\int_{-\infty}^{i}\right)^n-\sum_{i=-j+1}^{j+1}\left(\int_i^{\infty}\right)^n\right\}. \tag{16}$$

Equation (16) gives the mean range Ξ in a grouped sample in terms of powers of the probability integral of the parental population. For a normal distribution this is a particularly simple formula since such powers have already been calculated by L. H. C. Tippett (1925) and are conveniently tabulated in Table XXI of the *Tables for Statisticians and Biometricians*, Part II. A table of Ξ can, therefore, be easily computed by adding a few entries from Tippett's table and deducting the sum from the appropriate value of $2j+1$.

This has been done for samples of five, ten and twenty observations grouped in categories of breadth h (see table on p. 339). The parameter ξ denotes the

distance of the population mean $X = 0$ from the nearest group end-point. For given h the mean range Ξ in grouped samples is obviously a symmetrical periodic function of ξ with period h. The table has been extended to cover rather coarse grouping intervals ($h = 2{\cdot}2\sigma$) in order to illustrate the possible bias of range when estimated from frequency tables with as few as two or three categories. It is apparent that for small or moderate group intervals, say $h \leqslant \sigma$, the mean range is practically independent of h and ξ, so that no correction (corresponding to the well-known Sheppard's correction for the sample standard deviation) is required for the

Table of mean range in grouped samples drawn from a normal population having unit standard deviation

Size of sample $= n$. Width of group interval $= h$.
Distance of population mean to nearest group-end point $= \xi$.

h	ξ	$n=5$	$n=10$	$n=20$
0·2	0·0	2·32 593	3·07 751	3·73 495
0·6	0·0	2·32 593	3·07 750	3·73 500
	0·2	2·32 593	3·07 751	3·73 492
1·0	0·0	2·32 532	3·08 122	3·72 917
	0·2	2·32 574	3·07 865	3·73 317
	0·4	2·32 642	3·07 450	3·73 962
1·4	0·0	2·31 042	3·06 204	3·82 069
	0·2	2·31 626	3·06 787	3·78 826
	0·4	2·32 938	3·08 095	3·71 575
	0·6	2·33 990	3·09 143	3·65 796
1·8	0·0	2·29 227	2·90 539	3·67 974
	0·2	2·30 022	2·94 491	3·69 553
	0·4	2·32 023	3·04 621	3·73 085
	0·6	2·34 276	3·16 356	3·76 263
	0·8	2·35 734	3·24 140	3·77 838
2·2	0·0	2·36 011	2·77 080	3·27 509
	0·2	2·35 552	2·81 669	3·34 611
	0·4	2·34 221	2·94 307	3·53 896
	0·6	2·32 265	3·11 581	3·79 662
	0·8	2·30 255	3·28 173	4·03 847
	1·0	2·28 963	3·38 358	4·18 452

range. For $h = 0{\cdot}2\sigma$ the mean range in the grouped sample agrees with the theoretical ungrouped range to five places of decimals (see Table of Mean Range, Table XXII of *Tables for Statisticians and Biometricians*, Part II). For coarse grouping the correction becomes important but depends on ξ (as well as on h). For fixed h, as ξ varies from $-\frac{1}{2}h$ to $+\frac{1}{2}h$ the grouped mean range oscillates about the true mean range as a smooth single-period function. The reason for this is obvious. If ξ has a position such that the average positions of x_n and x_1 both happen to fall within the outside halves of two group intervals, then $\xi_n - \xi_1$ will

on the average, be smaller than $x_n - x_1$; vice versa, if the average positions of x_n and x_1 are in the inside halves of two group intervals, there will be a predominance of samples for which $\xi_n - \xi_1$ is larger than $x_n - x_1$.

Moreover, as h increases, the grouped mean range becomes a less reliable estimate, and it can be shown that for $h > \sigma$ the standard deviation of the random sampling distribution of the grouped range rapidly increases with h.

We are thus led to consider two problems; one is the elimination of the parameter ξ (or the dependence of the distribution on the position of the parental population mean); the other is to investigate more closely the random sampling distribution of the grouped range. Before dealing with these problems, however, we must first consider the distribution of the true range (range in the ungrouped sample).

4. The probability integral of the range in random samples

As before we denote by $x_1, \ldots, x_n$ the observations in a random sample drawn from a parental distribution $f(x)$ and arranged in ascending order of magnitude. The range in such a sample, defined as $w = x_n - x_1$, may be regarded as the limit of the grouped range $\xi_n - \xi_1$ as h, the group interval, tends to 0, i.e.

$$w = \lim_{h \to 0} \xi_n - \xi_1.$$

The probability integral of the range w, denoted by $P_n(W)$, is therefore the limit of $P(n, h, m-1, \xi)$, given by (3), as h tends to 0. To obtain this limit we write equation (3) as follows:

$$\begin{aligned} P(n, h, m-1, \xi) &= \sum_{i=-\infty}^{+\infty} \left(\int_{\xi+ih}^{\xi+(i+m)h} f(x)\,dx \right)^n - \left(\int_{\xi+(i+1)h}^{\xi+(i+m)h} f(x)\,dx \right)^n \\ &= \sum_{i=-\infty}^{+\infty} n \left(\int_{\bar{\xi}_i}^{\xi+(i+m)h} f(x)\,dx \right)^{n-1} f(\bar{\xi}_i)\,h, \end{aligned}$$

where $\bar{\xi}_i$ is some mean value between $\xi + ih$ and $\xi + (i+1)h$.

We now put $\qquad m = W/h \quad \text{or} \quad W = mh,$

and let h tend to 0, m to ∞, keeping W constant. We obtain without difficulty

$$\begin{aligned} P_n(W) &= \lim_{\substack{h \to 0 \\ m \to \infty}} P(n, h, m-1, \xi) \\ &= n \int_{-\infty}^{+\infty} f(\xi) \left(\int_{\xi}^{\xi+W} f(x)\,dx \right)^{n-1} d\xi, \end{aligned} \tag{17}$$

which is the required probability integral of the range. This integral may be compared with the expression for the distribution function of w which was given by A. T. McKay & E. S. Pearson (1933). It is easily verified that the function $\phi(w)$ given by these authors is the differential of $P(W)$. The expression for $P(W)$

is decidedly simpler than that for $\phi(w)$ which was used by Pearson for numerical work on this function. However, even $P(W)$ is of a complex character, and only in special cases is it possible to evaluate it analytically. For the rectangular distribution function ($f(x) \equiv 1$, $0 \leqslant x \leqslant 1$) this can be done easily.

5. THE PROBABILITY INTEGRAL OF THE RANGE IN SAMPLES FROM A NORMAL POPULATION

Of particular interest is the case where the parental population is normal. In this case we have

$$f(x) = z(x) = \frac{1}{\sqrt{(2\pi)}}\, e^{-\frac{1}{2}x^2}. \tag{18}$$

L. H. C. Tippett (1925), E. S. Pearson (1926, 1932) and A. T. McKay & E. S. Pearson (1933) have considered this problem and carried out extensive numerical work. The method adopted was to calculate correct values of the means and standard deviations of the distributions (as functions of n) and then, with the help of approximate values of β_1 and β_2 use as approximations to the unknown true distribution Pearson-type curves fitted by the method of moments. The numerical results, although they have been successfully tested by experimental sampling, have, of course, an unknown accuracy. It is therefore desirable to find a method which produces $P_n(W)$ to known and sufficient accuracy.

We have
$$P_n(W) = n\int_{-\infty}^{+\infty} z(\xi)\left(\int_{\xi}^{\xi+W} z(x)\,dx\right)^{n-1} d\xi$$
$$= n\int_{-\infty}^{-\frac{1}{2}W} + n\int_{-\frac{1}{2}W}^{+\infty} = I_1 + I_2 \quad \text{(say)}.$$

Writing
$$\eta = -(\xi + W), \quad -\xi = \eta + W,$$

we obtain
$$P_n(W) = n\int_{-\frac{1}{2}W}^{\infty} z(-\eta - W)\left(\int_{-(\eta+W)}^{-\eta} z(x)\,dx\right)^{n-1} d\eta + I_2.$$

Using the symmetry of $z(x)$ we may write

$$P_n(W) = n\int_{-\frac{1}{2}W}^{\infty} z(\eta + W)\left(\int_{\eta}^{\eta+W} z(x)\,dx\right)^{n-1} d\eta + I_2,$$

and writing ξ as a variable of integration in place of η we find

$$P_n(W) = n\int_{-\frac{1}{2}W}^{\infty} [z(\xi + W) + z(\xi)]\left(\int_{\xi}^{\xi+W} z(x)\,dx\right)^{n-1} d\xi,$$

or
$$P_n(W) = -n\int_{-\frac{1}{2}W}^{\infty} [z(\xi + W) - z(\xi)]\left(\int_{\xi}^{\xi+W} z(x)\,dx\right)^{n-1} d\xi$$
$$+ 2n\int_{-\frac{1}{2}W}^{\infty} z(\xi + W)\left(\int_{\xi}^{\xi+W} z(x)\,dx\right)^{n-1} d\xi. \tag{19}$$

Integrating the first integral and introducing $u = \xi + W$ in the second integral we finally obtain

$$P_n(W) = \left(\int_{-\frac{1}{2}W}^{+\frac{1}{2}W} z(x)\,dx\right)^n + 2n\int_{\frac{1}{2}W}^{\infty} z(u)\left(\int_{u-W}^{u} z(x)\,dx\right)^{n-1} du. \qquad (20)$$

For large values of W this is an approximate solution of the problem since the second term in (20) is small, so that the first term

$$\left(\int_{-\frac{1}{2}W}^{+\frac{1}{2}W} z(x)\,dx\right)^n$$

gives a fair approximation to $P_n(W)$. This expression denotes the chance of observing samples with observations all lying between $-\frac{1}{2}W$ and $+\frac{1}{2}W$; all these samples have a range smaller than or equal to W. For large W it is these samples which constitute an ever-increasing proportion of the total number of samples with range $\leqslant W$.

The second term in (20), which is always positive, takes into account all those samples which are not contained in the interval $-\frac{1}{2}W$ to $+\frac{1}{2}W$. This term cannot be ignored if high accuracy is required and if W is small or moderate. Nevertheless, the work involved in the numerical integration has been considerably reduced, for the range of integration is now from $+\frac{1}{2}W$ to $+\infty$.

The numerical integration of

$$\int_{\frac{1}{2}W}^{\infty} z(u)\left(\int_{u-W}^{u} z(x)\,dx\right)^{n-1} du$$

is best carried out simultaneously for values of n forming an arithmetical progression. For fixed u and W, the integrand is then given by the terms of a geometrical progression with, say,

$$z(u)\left(\int_{u-W}^{u} z(x)\,dx\right)^2$$

as first term and

$$\left(\int_{u-W}^{u} z(x)\,dx\right)$$

as common ratio. Such a geometrical progression can be produced automatically by certain modern calculating machines. This forms the basic idea of the actual computation of the probability integral which is described in detail in another paper (pp. 309–10 above).

6. The range in randomly grouped samples

The results of sections (2) and (3) on the effect of grouping on the distribution of range depend on the parameter ξ, which denotes the origin of the equidistant set of group end-points

$$\xi + ih \quad i = 0, 1, 2, \ldots,$$
$$-1, -2, \ldots,$$

given by equation (1). In practice, however, all we know is the category breadth, h.

We then select the actual group end-points $\xi+ih$ from considerations which are, by necessity, independent of the position of the mean X of the parental population $f(x)$, because this position will generally be unknown. One of our group end-points, however, is bound to fall into the interval

$$X-\tfrac{1}{2}h \quad \text{to} \quad X+\tfrac{1}{2}h,$$

wherever this interval may lie. Now, since the origin ξ in our system of group end-points is wholly arbitrary we may assume that for given h we have the inequality

$$X-\tfrac{1}{2}h \leqslant \xi \leqslant X+\tfrac{1}{2}h.$$

The fact that our group end-points (and therefore their origin ξ) are chosen independently of X has now to be expressed in mathematical terms. This is done by assuming that we are dealing with a population of values of ξ (being the origins of corresponding systems of group end-points) which are rectangularly distributed in the interval $X-\frac{1}{2}h \leqslant \xi \leqslant X+\frac{1}{2}h$. This condition is exactly fulfilled where grouping has been introduced through rounding off of data (example (*a*) on p. 335), and it is often an appropriate assumption in the other examples, as in many other cases of grouping which occur in practice.*

In order to derive the distribution of range in samples randomly grouped in the above sense we have to return to section (2). In this section we derived the probability $P(n,h,m-1,\xi)$, giving the chance that the difference between the centre points ξ_n and ξ_1 of the highest and lowest category covered by a sample of n items is at most $(m-1)h$, where h is the constant category breadth and group end-points are given by (1). The frequency distribution of $\xi_n-\xi_1$ which we may denote by $\phi(n,h,m,\xi)$ is then given by

$$\phi(n,h,m,\xi) = P(n,h,m,\xi)-P(n,h,m-1,\xi),$$

and represents the chance that $\xi_n-\xi_1$ is exactly mh, given a particular value of ξ. The corresponding frequency distribution for random grouping may be denoted by $\phi(n,h,m)$. To derive it we may apply Bayes's Theorem and obtain

$$\phi(n,h,m) = \frac{1}{h}\int_{X-\frac{1}{2}h}^{X+\frac{1}{2}h} \phi(n,h,m,\xi)\,d\xi.$$

The resulting cumulative probability may therefore be defined by

$$P(n,h,m) = \sum_{j=1}^{m} \phi(n,h,j) = \frac{1}{h}\int_{X-\frac{1}{2}h}^{X+\frac{1}{2}h} \sum_{j=1}^{m} \phi(n,h,j,\xi)\,d\xi,$$

which yields the corresponding relation for the cumulative probabilities

$$P(n,h,m) = \frac{1}{h}\int_{X-\frac{1}{2}h}^{X+\frac{1}{2}h} P(n,h,m,\xi)\,d\xi.$$

* In certain cases, when grouping is very coarse, it may be advantageous to use an estimate $\bar{x}$ of the population mean X (either dependent or independent of the sample whose range is considered). The increase in information is akin to that given by an ancillary statistic in estimation theory. However, from the results in § 3 it would appear that little information is gained where the grouping interval is small or moderate.

Substituting, now, the expression (3) for $P(n, h, m, \xi)$ we obtain

$$P(n,h,m) = \frac{1}{h}\sum_{i=-\infty}^{+\infty}\int_{X-\frac{1}{2}h}^{X+\frac{1}{2}h}\left\{\left(\int_{\xi+ih}^{\xi+(i+m+1)h} f(x)\,dx\right)^n - \left(\int_{\xi+(i+1)h}^{\xi+(i+m+1)h} f(x)\,dx\right)^n\right\}d\xi$$

$$= \frac{1}{h}\int_{-\infty}^{+\infty}\left\{\left(\int_{\xi}^{\xi+(m+1)h} f(x)\,dx\right)^n - \left(\int_{\xi}^{\xi+mh} f(x)\,dx\right)^n\right\}d\xi. \qquad (21)$$

This formula may be reduced to a simpler form in which its relation to the probability integral of the true range (17) becomes apparent. We introduce the second integral

$$F_n(W) = \int_0^W P_n(w)\,dw$$

$$= \int_0^W n\int_{-\infty}^{+\infty} f(\xi)\left(\int_{\xi}^{\xi+w} f(x)\,dx\right)^{n-1} d\xi\,dw. \qquad (22)$$

The first integration is with regard to w and the integrand is an integral with regard to ξ. If, now, in this latter integral $\eta = \xi + w$ is used as variable of integration in place of ξ we have

$$F_n(W) = \int_0^W n\int_{-\infty}^{+\infty} f(\eta - w)\left(\int_{\eta-w}^{\eta} f(x)\,dx\right)^{n-1} d\eta\,dw.$$

We now note that the integrand may be written as a differential with regard to w. Thus, interchanging the order of integration we obtain

$$F_n(W) = \int_{-\infty}^{+\infty}\int_0^W \frac{d}{dw}\left\{\int_{\eta-w}^{\eta} f(x)\,dx\right\}^n dw\,d\eta$$

$$= \int_{-\infty}^{+\infty}\left\{\int_{\eta-W}^{\eta} f(x)\,dx\right\}^n d\eta, \qquad (23)$$

thus eliminating integration with regard to w. The surviving variable of integration η may now be replaced by $\xi = \eta - W$ so that we reach the final result

$$F_n(W) = \int_{-\infty}^{+\infty}\left(\int_{\xi}^{\xi+W} f(x)\,dx\right)^n d\xi. \qquad (24)$$

We now observe that this integral is identical with the one occurring in the expression (21) for $P(n, h, m)$, and we note the relation

$$P(n,h,m) = \frac{1}{h}\{F_n(\overline{m+1}\,h) - F_n(mh)\}$$

$$= \frac{1}{h}\int_{mh}^{(m+1)h} P_n(w)\,dw. \qquad (25)$$

This simple formula makes it possible to obtain the effect of random grouping on the probability integral of range $P_n(W)$. In particular, for normal samples for which $P_n(W)$ has been tabulated at the fine interval of $\Delta w = 0{\cdot}05$, the second integral $F_n(W)$ is easily obtained numerically by summation of the tabular entries in the table on pp. 302–7 above.

For $h \to 0$, $m \to \infty$, $mh \to W$, we note that, as expected, $P(n, h, m) \rightrightarrows P_n(W)$; that is, as grouping becomes finer and finer the probability integral of the grouped range tends to that of the true range. What is not quite obvious, however, is the identity of mean grouped range and mean true range, no matter how large the breadth, h, of the randomly placed groups. This point we shall now examine.

If we denote by $w(n, h)$ the mean range in samples of n items classified in groups of breadth h randomly placed, we have by definition

$$w(n,h) = \sum_{m=0}^{\infty} \phi(n,h,m)\, mh = \sum_{m=0}^{\infty} \{P(n,h,m) - P(n,h,m-1)\}\, mh. \qquad (26)$$

We now introduce central differences of the function $F_n(W)$ and use the notation

$$\Delta'_{m+\frac{1}{2}} = F_n(\overline{m+1}h) - F_n(mh), \quad \Delta''_m = F_n(\overline{m+1}h) - 2F_n(mh) + F_n(\overline{m-1}h), \qquad (27)$$

so that we obtain from (25), (26) and (27)

$$w(n,h) = \sum_{m=1}^{\infty} \Delta''_m\, m.$$

We may now write

$$\begin{aligned} w(n,h) &= \lim_{M\to\infty} \sum_{m=1}^{M} \Delta''_m\, m \\ &= \lim_{M\to\infty} \left\{ M\Delta'_{M+\frac{1}{2}} - \sum_{m=0}^{M-1} \Delta'_{m+\frac{1}{2}} \right\} \\ &= \lim_{M\to\infty} \{M\Delta'_{M+\frac{1}{2}} - F_n(Mh)\}. \end{aligned} \qquad (28)$$

On the other hand, we have for the mean true range ($\overline{w}_n$ say)

$$\overline{w}_n = \int_0^{\infty} w f_n(w)\, dw,$$

where $f_n(w) = \dfrac{d}{dw} P_n(w)$ is the distribution function of the true range. We, therefore, have

$$\begin{aligned} \overline{w}_n &= \lim_{M\to\infty} \int_0^{Mh} w f_n(w)\, dw \\ &= \lim_{M\to\infty} \left\{ MhP_n(Mh) - \int_0^{Mh} P_n(w)\, dw \right\}. \end{aligned} \qquad (29)$$

On taking the difference of (28) and (29) we see that

$$w(n,h) - \overline{w}_n = \lim_{M\to\infty} Mh \left\{ \frac{1}{h}\Delta'_{M+\frac{1}{2}} - P_n(Mh) \right\}.$$

Now $$\frac{1}{h}\Delta'_{M+\frac{1}{2}} = \frac{1}{h}\int_{Mh}^{(M+1)h} P_n(w)\, dw = P_n(Mh) + f_n(w^*)\, h^*,$$

where $$Mh < w^* < (M+1)h \quad \text{and} \quad h^* \leqslant h,$$

so that $$|\, w(n,h) - \overline{w}_n \,| \leqslant \lim_{M\to\infty} f_n(w^*)\, Mh^2.$$

From this relation it is obvious that $w(n,h) = \overline{w}_n$, since $f_n(w)$ has contact of at least second order as $w \to \infty$.*

We have proved, therefore, that if grouping is random the expectation† of the mean range $w(n,h)$ (mean grouped range) is identical with the true mean range $\overline{w}_n$ so that no bias is introduced through random grouping, no matter how large the grouping interval h. However, if we wish to use $\xi_n - \xi_1$ as an estimate of w_n this estimate, although unbiased in the sense defined, becomes less and less reliable as h increases. This is borne out by its random sampling distribution or its probability integral $P(n,h,m)$ given by (25). For normal samples it is an easy matter to tabulate $P(n,h,m)$ from the table of $P_n(W)$ (pp. 302–7) and thus to follow up the numerical increase of its standard deviation as h increases. However, to cover the case of a general parental distribution $f(x)$, we shall derive an analytical formula for the variance of $\xi_n - \xi_1$ from which approximate numerical results are easily obtained.

In order to obtain this formula we consider the second moment of $\xi_n - \xi_1$ which we may denote by $\mu_2(n,h)$. We have by definition and from equations (25), (26) and (27)

$$\begin{aligned}\mu_2(n,h) &= \lim_{M\to\infty} \sum_{m=0}^{M} m^2h^2 \frac{1}{h} \Delta''_m \\ &= \lim_{M\to\infty} \sum_{m=0}^{M} hm(m+\tfrac{1}{2})\,\Delta''_m - \tfrac{1}{2}h\overline{w}_n. \end{aligned} \tag{30}$$

Now we may write

$$\begin{aligned}\sum_{m=0}^{M} m(m+\tfrac{1}{2})\,\Delta''_m &= \sum_{m=0}^{M} m\{(m+\tfrac{1}{2})\,\Delta'_{m+\frac{1}{2}} - (m-\tfrac{1}{2})\,\Delta'_{m-\frac{1}{2}}\} - \sum_{m=0}^{M} m\Delta'_{m-\frac{1}{2}} \\ &= M(M+\tfrac{1}{2})\,\Delta'_{M+\frac{1}{2}} - \sum_{m=0}^{M-1} (m+\tfrac{1}{2})\,\Delta'_{m+\frac{1}{2}} - MF_n(Mh) + \sum_{m=0}^{M-1} F_n(mh) \\ &= M(M+\tfrac{1}{2})\,\Delta'_{M+\frac{1}{2}} - (2M-\tfrac{1}{2})\,F_n(Mh) + 2\sum_{m=0}^{M-1} F_n(mh). \end{aligned} \tag{31}$$

Using equation (31) we obtain for the second moment (30)

$$\mu_2(n,h) = \lim_{M\to\infty} \left\{(Mh)^2\left(1+\frac{1}{2M}\right) - 2MhF_n(Mh)\left(1-\frac{1}{4M}\right) + 2h\sum_{m=0}^{M-1} F_n(mh) - \frac{h}{2}\overline{w}_n\right\}. \tag{32}$$

This formula enables us to compare $\mu_2(n,h)$ with $\mu_2(n)$, the second moment of the distribution of the true range. We have by definition

$$\mu_2(n) = \lim_{M\to\infty} \int_0^{Mh} f_n(w)\, w^2\, dw,$$

* It can be proved that the order of contact of $f_n(w)$ is the same as that of the parental distribution $f(x)$.

† If repeated samples were drawn from the same population and the same grouping system used in each case, the mean grouped range would be biased by an unknown amount. But in repeated experience with different populations the expectation of this bias is zero.

which we transform by two partial integrations into the equation

$$\mu_2(n) = \lim_{M\to\infty}\left\{(Mh)^2 P_n(Mh) - 2(Mh)F_n(Mh) + 2\int_0^{Mh} F_n(w)\,dw\right\}. \qquad (33)$$

Taking the difference of (32) and (33) we obtain

$$\mu_2(n,h) - \mu_2(n) = \lim_{M\to\infty}\left\{\frac{Mh}{2}h - \frac{hF_n(Mh)}{2} - \frac{h}{2}\overline{w}_n\right\}$$
$$+ \lim_{M\to\infty}\left\{2h\sum_{m=0}^{M-1} F_n(mh) + hF_n(Mh) - 2\int_0^{Mh} F_n(w)\,dw\right\}. \qquad (34)$$

Since the expected mean value of the grouped range is the same as for the true range, the difference in second moments about zero equals the difference in variances. The first term on the right-hand side of equation (34) is obviously 0 (see equation (29)), whilst the second term is best evaluated with the help of Gregory's formula for numerical integration (see e.g. L. J. Comrie, 1936, p. 809). Using this formula we can express the difference between the integral and the finite sum in equation (34) in terms of the differences of the integrand $F_n(w)$ at the two ends of the range of integration. We obtain

$$\mu_2(n,h) - \mu_2(n) = \lim_{M\to\infty}\left\{\frac{2h}{12}\{\Delta'_{M-\frac{1}{2}} - \Delta'_{\frac{1}{2}}\} + \frac{2h}{24}\{\Delta''_{M-1} + \Delta''_1\} + \ldots\right\}$$
$$= \tfrac{1}{6}h\,(h - \Delta'_{\frac{1}{2}}) + \frac{1}{12}h\Delta''_1 + \ldots, \qquad (35)$$

provided the Gregory expansion is convergent.* $\Delta'_{\frac{1}{2}}$, Δ''_1, $\Delta'''_{\frac{1}{2}}$, ... are advancing differences of the function $F_n(w)$ at $w = 0$.

Equation (35) yields the desired formula for the second moment of $\xi_n - \xi_1$. For most parental distributions the resulting probability integral of the range will be practically 0 for a certain range in the neighbourhood of the origin (see for instance the behaviour of $P_n(W)$ from a normal parent given in the table on pp. 302–7). For such parents and for moderate h we have

$$\Delta'_{\frac{1}{2}} \cong \Delta''_1 \cong \Delta'''_{\frac{1}{2}} \cong \ldots \cong 0,$$

so that

$$\mu_2(n,h) - \mu_2(n) \cong \tfrac{1}{6}h^2. \qquad (36)$$

For small or moderate values of h, therefore, the increase in variance of the grouped range is given approximately (and for most parents to a high degree of accuracy) by $\frac{1}{6}h^2$. This increase is double the amount given by the well-known Sheppards correction of $\frac{1}{12}h^2$. Indeed, had we grouped a *sample of true ranges* w in fixed categories of breadth h, the resulting second moment of the grouped distribution of w would have an expectation which is $\frac{1}{12}h^2$ in excess of the second moment of the true range. With *random grouping* of the *original sample* (as it

* This condition is as a rule fulfilled for values of h which do not exceed the standard deviation of the parental distribution $f(x)$.

has been defined above in accordance with common practice of grouping) an additional uncertainty is introduced by using a new, randomly selected, set of group intervals each time a new grouped range is determined. This additional uncertainty has been proved roughly to double the excess of the variance and the result is an increase of $\frac{1}{6}h^2$ over the variance of the true range.

I wish to acknowledge with gratitude the helpful suggestions and criticisms made by Drs J. Wishart and J. O. Irwin and by Professor E. S. Pearson at various stages of this investigation.

REFERENCES

COMRIE, L. J. (1936). Interpolation and Allied Tables. H.M.S.O.
MCKAY, A. T. (1935). *Biometrika*, **27**, 466–71.
MCKAY, A. T. & PEARSON, E. S. (1933). *Biometrika*, **25**, 415–20.
PEARSON, E. S. (1926). *Biometrika*, **18**, 173–94.
—— (1932). *Biometrika*, **24**, 404–17.
PEARSON, E. S. & HAINES, J. (1935). *J.R. Statist. Soc.* Suppl. **2**, 83–98.
TILDESLEY, M. L. (1940). *Man*, **40**, 180–9.
TIPPETT, L. H. C. (1925). *Biometrika*, **17**, 364–87.

MISCELLANEA

(i) The Second Yearbook of Research and Statistical Methodology Books and Reviews. Edited by OSCAR KRISEN BUROS. The Gryphon Press, Highland Park, New Jersey, 1941. $5.

This is a second and much enlarged issue of a volume published in 1938. It contains nearly seventeen hundred review excerpts on 346 statistical and allied books (in the English language only), extracted from 283 different journals. The editor has attempted with considerable success to cover the whole field of statistical and probability theory, as well as their applications in every possible direction. He has also included reviews of a number of books on the general history of science, on scientific method and on the social relations of science on the ground that they are—or should be—of general interest to scientific workers in every special field. Included in this category are books such as J. D. Bernal's *The Social Function of Science*, J. G. Crowther's *The Social Relations of Science* and J. B. S. Haldane's *The Marxist Philosophy and the Sciences*.

This large volume of several hundred pages has been admirably produced and arranged. It is intended to publish a fresh volume every two years containing reviews that have appeared in the interval. The Preface sets out a variety of reasons which, in the Editor's opinion, justify the present venture and even its enlargement in the future if sufficient support is forthcoming; at the same time frank expressions of opinions are asked for from readers and reviewers.

The objectives of the *Yearbook* as set out may be classed under four general heads:

(1) To help students, teachers and librarians to select text-books with greater discrimination and to point out to them the weak and strong points of particular books.

(2) To indicate the width of the subject of statistics and the many fields in which it is applied.

(3) To make students and teachers aware of the inadequacy of much that is now presented in text books and classes; to discourage the publication of books written by persons ignorant of the latest developments in their subject.

(4) To improve the quality of reviews by encouraging editors and reviewers alike to take their responsibilities more seriously.

With the last three objectives it is hardly possible to quarrel, and it is likely that the wide circulation of this volume would provide one of the most direct methods of attaining these ends. The first objective is, however, presumably the most important, and there are bound to be differences of opinion on the probable success of the book in this direction. In the ordinary event the teacher will no doubt be made aware of new books in the field with which he is concerned by reading the notices in one or two journals specially devoted to his subject. Having obtained a suggestion of a likely book he must surely get hold of it and determine by reading it himself whether it is suitable and abreast of the latest developments. If he is not competent to do this, but must base his decision on the advice of 6–10 reviewers, it seems doubtful whether he should be teaching the subject at all.

After reading through the reviews on some dozen books contained in the present *Yearbook*, I am inclined to the following conclusions. Regarding books of outstanding but perhaps rather controversial character, as those of Harold Jeffreys and Richard von Mises, the reader will certainly gain a useful impression from the collected reviews. This is partly because in such cases the standing of the reviewers is high and their reviews interesting and fairly written, even if critical. But in the case of the more elementary text book, the position is rather different. Quite often the opinions expressed are diametrically opposite. In cases where I knew nothing of the book or its author I found myself inevitably forced to form an opinion from my own personal knowledge of the experience, the special interests

and even the character of the reviewer. Such inside information will generally not be possessed by the College instructor and certainly not by the student. We cannot, I think, escape the conclusion that the teacher who has to select a text book for his students must be competent to decide on its merits himself and, if he is not, he may be only confused by the varied opinions contained in the *Yearbook*.

Three future directions in which the volume might be enlarged are contemplated:

(1) The inclusion of reviews of foreign language (i.e. not English) books.

(2) The addition of a section devoted to non-critical abstracts of periodical literature on research and statistical methods.

(3) The publication of original criticisms by one or more persons (according to the importance and controversial nature) of articles and papers in the periodical literature.

The first addition is clearly desirable; the publication of translations of reviews in foreign journals of our own American and British books would probably be useful too; it would help us to see ourselves as others see us. With regard to the second and third proposals, the great difficulty is of course to secure the services of sufficiently competent abstractors or critics for so large an undertaking. If, quoting the Editor, the statistical student and teacher are to be kept 'abreast of modern developments in statistical theory'; to be warned 'to ignore much of the literature which either presents nothing new or presents inefficient or incorrect methods of statistical analysis'; to be told what are 'sloppy, valueless, and erroneous articles' and what are 'well-written, significant contributions', it is clear that a very great responsibility will lie on the Editor of the *Yearbook* and his collaborators. As Prof. Buros indicates, the organizing and editing of such a comprehensive service would need the support of a foundation interested in fostering the advancement of research. Indeed, to avoid duplication the organization must be built up on an international basis, possibly in collaboration with such bodies as the American Statistical Association and the Royal Statistical Society between whose representatives some discussion of a similar project took place a few years ago.

E. S. P.

BIOMETRIKA